Controlling chaotic dynamics of periodically forced spheroids in simple shear flow: Results for an example of a potential application

C V ANIL KUMAR and T R RAMAMOHAN

Computational Materials Science, Regional Research Laboratory (CSIR), Thiruvananthapuram 695 019, India
e-mail: ram@csrrltrd.ren.nic.in

MS received 17 December 1996; revised 11 April 1997

Abstract. Recently, we studied the technologically important problem of periodically forced spheroids in simple shear flow and demonstrated the existence of chaotic parametric regimes. Our results indicated a strong dependence of the solutions obtained on the aspect ratio of the spheroids, which can be used to separate particles from a suspension. In this paper we demonstrate that controlling the chaotic dynamics of periodically forced particles by a suitably engineered novel control technique, which needs little information about the system and is easy to implement, leads to the possibility of better separation. Utilizing the flexibility of controlling chaotic dynamics in a desired orbit irrespective of initial state, we show that it is theoretically possible to separate particles much more efficiently than otherwise from a suspension of particles having different shapes but similar sizes especially for particles of aspect ratio $r_e > 1.0$. The strong dependence of the controlled orbit on the aspect ratio of the particles may have many applications such as in the development of computer-controlled intelligent rheology. The results suggest that control of chaos as discussed in this work may also have many applications.

Keywords. Periodically forced spheroids; chaos; control; particle separation.

1. Introduction

There has been appreciable interest in the dynamics of suspensions of small dipolar particles stabilized by surface active agents in a variety of linear flows under the effect of alternating or rotating external fields. This problem has been studied by a number of authors and a number of practical applications such as magnetofluidization (Buevich *et al*

A list of symbols is given at the end of the paper

1984), magnetostriction of ferromagnetic particle suspensions (Ignatenko *et al* 1984), the growth of single crystals from a melt by Czochralski's method (Zibol'd *et al* 1986), rheological properties of ferromagnetic colloids (Tsebers 1986) and the characterization of magnetorheological suspensions (Cebers 1993; Petrikevich & Raikher 1984; Kashevskii 1986; Shul'man *et al* 1986) have been discussed. Some of the above authors consider rotatory external fields which can be taken as a superposition of two orthogonal alternating fields. As can be seen from the relevant literature cited above, the problem discussed in this paper is of interest in a number of technological applications. For a brief literature review of the importance, relevance and novelty of this problem, the reader is referred to our earlier paper (Kumar *et al* 1995) and the references therein. In a preliminary analysis of this problem we have demonstrated (Kumar *et al* 1995) the existence of chaotic dynamics and a strong dependence on the aspect ratio of the particles on the solution of the problem in the chaotic regime. This strong dependence was considered to demonstrate the potential of this problem to develop techniques for effective separation of particles by aspect ratio to yield well-characterized suspensions for testing theories of the above processes and for developing suitable theories for suspensions of particles of various aspect ratios and sizes. It is also of interest to study the possibility of controlling the chaos present in the dynamics of periodically forced particles with a view to develop potential applications such as control of magnetorheological suspensions, magnetostriction of ferromagnetic particle suspensions, magnetodielectric material suspensions, ferromagnetic colloidal suspensions and the possibility of computer-controlled intelligent rheology.

In order to separate out the effects of particle size and shape on the properties of the suspension, it would be desirable to obtain suspensions of particles having a narrow distribution of shapes and sizes. A narrow distribution of sizes can be obtained through appropriate separation techniques such as filtration. However, we are not aware of any effective technique to separate particles having the same size but different shapes.

There has been considerable interest recently in the use of chaos and controlling chaos as a tool to develop new processing technologies (Zumbrunnen *et al* 1996). Practically speaking, chaos may be desirable or undesirable depending on the application desired. In many practical situations, in order to improve system performance, the chaotic system must be controlled to a periodic orbit or to a steady state. Thus the technique of controlling chaotic systems has received increased attention in the recent literature. In a physical apparatus, one can imagine that the system dynamics is to be changed in some way so that improved performance is obtained. One can achieve this by making appropriate changes in the system to achieve the desired objective. Similar or otherwise unattainable objectives can be attained by operating the system in a chaotic regime and then applying suitable control of chaos techniques. Since a chaotic attractor is a closure of unstable periodic orbits of different periods, any of these orbits can be stabilized to attain the desired objective by suitable control algorithms.

The existence of chaotic parametric regimes in the problem considered in this work and the possibility of control of its chaotic behaviour (Ott *et al* 1990) allows the possibility of many applications including the development of computer controlled intelligent rheology. Another possible application of controlling the chaos present in this problem is further fine-tuning of the separation of the particles by aspect ratio than is possible with other methods. Since the chaotic phase space is an unlimited reservoir of periodic solutions

(Ott *et al* 1990), a particle of a particular aspect ratio can be forced to oscillate in any one of the orbits and can be picked up from there effectively. Though the results presented herein represent a preliminary analysis of the problem considered in dilute suspensions, the existence of chaos and hence its controllability and the strong dependence of the results on the aspect ratio of the controlled orbit suggest that these results may have many potential applications. In this paper, as an example, we focus on the more efficient separation of particles by aspect ratio through control of chaos.

2. Theory

The general equation governing the dynamics of suspensions of small dipolar spheroids of finite aspect ratio in simple shear flow under the effect of externally induced periodic fields, as developed by Kumar *et al* (1995), leads to

$$\begin{aligned}
\dot{\mathbf{u}}_1 &= Pu_2(1-2u_1^2) + Qu_2 + Ra_1\cos(\omega t),\\
\dot{\mathbf{u}}_2 &= Pu_1(1-2u_2^2) - Qu_1 + Ra_2\cos(\omega t),\\
\dot{\mathbf{u}}_3 &= -2Pu_1u_2u_3 + Ra_3\cos(\omega t), \qquad (1)\\
\text{where,}\quad a_1 &= u_3(u_3k_1 - u_1k_3) - u_2(u_1k_2 - u_2k_1),\\
a_2 &= u_1(u_1k_2 - u_2k_1) - u_3(u_2k_3 - u_3k_2),\\
a_3 &= u_2(u_2k_3 - u_3k_2) - u_1(u_3k_1 - u_1k_3),\\
P &= \frac{\pi(r_e^2-1)}{r_e},\quad Q = \frac{\pi(r_e^2+1)}{r_e},\quad R = \frac{3r_e^3[(2r_e^2-1)\beta - 1]}{8(r_e^2-1)},\\
\beta &= \frac{\ln[r_e + (r_e^2-1)^{1/2}]}{r_e(r_e^2-1)^{1/2}} \text{ or } \frac{\cos^{-1}(r_e)}{r_e(1-r_e^2)^{1/2}},
\end{aligned}$$

with β depending on the axis ratio r_e being > 1 or < 1, $\dot{\mathbf{u}}$ is the rate of change of the orientation vector; r_e is the particle axis ratio of the spheroid defined as $r_e = a/b$, where a = polar radius and b = equatorial radius; u_1, u_2, u_3 denote the x, y, z components of the unit vector $\mathbf{u}$ along the axis of symmetry which determines the orientation of the spheroids, respectively; k_1, k_2 and k_3 are the x, y, z components of the external force $\mathbf{k}$; ω is the frequency of the driver and t is the time.

As can be seen, the equations depend on the aspect ratio of the particle and also depend on the magnitude and frequency of the external force field. The maximum scaled value of ω, corresponding to the limit where the quasistatic Stokes' equations are valid corresponds to $J = 2\pi(r_e + r_e^{-1})$. The three equations listed above can be reduced to a system of two nonlinear nonautonomous ordinary differential equations (ODEs), since u_1, u_2 and u_3 evolve with time, from the given initial condition $|\mathbf{u}| = 1$, the magnitude of the vector $\mathbf{u}$ remains at unity. Thus, (1) can be converted into equations for θ and ϕ in spherical co-ordinates as given by them, where θ is the azimuthal angle and ϕ is the polar angle corresponding to a given vector $\mathbf{u}$ as

$$\begin{aligned}
\frac{\mathrm{d}\theta}{\mathrm{d}t} &= \frac{P}{2}\sin 2\theta \sin 2\phi\\
&\quad + R[k_1\cos\theta\cos\phi + k_2\cos\theta\sin\phi - k_3\sin\theta]\cos(\omega t),
\end{aligned}$$

$$\frac{d\phi}{dt} = P\cos 2\phi - Q + R\left[\frac{-k_1 \sin\phi + k_2 \cos\phi}{\sin\theta}\right]\cos(\omega t), \qquad (2)$$

$$|\theta|, |\phi| \leq \pi/2.$$

We note that the equation for $\dot{\phi}$ has a singularity at $\theta = 0$. The singularity of $\dot{\phi}$ can be removed by taking $\phi_1 = \phi \sin\theta$. The equations for $\dot{\theta}$ and $\dot{\phi}$ when integrated numerically, did not lead to any problems with the singularity at $\theta = 0$ and we obtained the same results as that obtained by the other set of equations. For a more detailed discussion of the method of solution of (2) the reader is referred to Kumar *et al* (1995). In this paper we suggest, as an example of the importance of chaos control, a technique to separate particles having the same size but different shapes much more efficiently than was demonstrated by them.

The fact that a chaotic solution eliminates the possibility of long-term prediction of system behaviour induced many reports in the literature of either quenching chaos or controlling chaos (Ott *et al* 1990; Ditto *et al* 1995). Since chaotic attractors have embedded within them a dense set of unstable periodic orbits, any one of the unstable periodic orbits can be stabilized to obtain otherwise unattainable system behaviour. The essential idea is that a chaotic system explores a relatively large region of state space and the system can be brought to a desired stable state to improve the performance of the separation technique by a suitable control algorithm. The first method (OGY – acronym formed from the authors' names) of control of chaos proposed by Ott *et al* (1990) generated appreciable interest in the literature of chaos. Thereafter, a large number of algorithms for controlling chaos have been reported in the literature (Christini & Collins 1995; Rhode *et al* 1995). Broadly speaking there are two classes of algorithms for controlling chaos, namely (1) *Feedback methods* (Ott *et al* 1990) and (2) *Non-feedback methods* (Güémez *et al* 1994). The first method needs appreciable information about system behaviour but is comparatively simple to implement experimentally. The second method may be difficult to implement experimentally.

The algorithm suggested in this work needs very little information about system behaviour but is rather easy to implement experimentally. This algorithm can be classified as a *non-feedback control method*. Rajasekar & Lakshmanan (1992, 1993) have investigated the applicability and effectiveness of various approaches of controlling chaos in the BVP oscillator. Suppression of chaos by periodic parametric perturbations in an experimental set-up of a Duffing oscillator is also reported in the literature by Fronzoni *et al* (1991). In the method proposed in this paper, the control parameter is perturbed for one period at fixed intervals of every integral multiple of the fundamental period. In the methods considered earlier in the literature (Lakshmanan & Murali 1996), the parameter was perturbed continuously rather than at fixed intervals. To implement the control strategy reported in this paper, we apply a constant external force in addition to the periodic force for the duration of one period after every $n-1$ periods as shown in figure 1, where length of one period (T) is equal to length of the fundamental period of the periodic force. If the system is perturbed after $n-1$ periods by applying an additional constant force, the system is found to be stabilized in a periodic orbit with a period equal to nT or with a period equal to an integral multiple of nT.

For example, if a represents the magnitude of the additional constant force, then the control can be implemented by setting $a = 0$ if $j \not\equiv 0 \bmod(n)$ and $a \neq 0$ if $j \equiv 0 \bmod(n)$ where j represents iteration number and the time interval between two successive

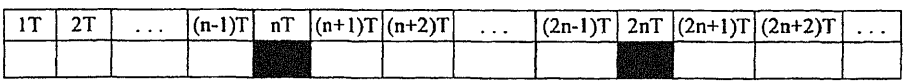


Figure 1. The schematic representation of the control strategy presented in this paper, where control is active during the period corresponding to the shaded area and not active during the period corresponding to the blank area.

erations is equal to the fundamental period T. That is, the system evolves without any odification for periods which are not multiples of n and with modification for periods hich are multiples of n. A schematic representation of the control strategy is given in gure 1. The chioce of n depends on the period-m solution we want to stabilise. The integer can be either m or a divisor of m depending on the value of a and the choice of n.

To incorporate the above idea of controlling the dynamics of the system, we modify 2) governing the dynamics of the system by introducing an additional constant force in ddition to the periodic force along the direction of the periodic force. Let k_1', k_2' and k_3' e the x, y, z components of the additional constant force, $\mathbf{k}'$. After scaling all quantities ppearing in the equations similarly to Kumar *et al* (1995), (2) can be written as

$$\begin{aligned}\frac{\mathrm{d}\theta}{\mathrm{d}t} &= \frac{P}{2}\sin 2\theta \sin 2\phi + R[(\cos\theta\cos\phi\ k_1 + \cos\theta\sin\phi\ k_2 - \sin\theta\ k_3)\\ &\quad \times \cos(\omega t)] + R[(\cos\theta\cos\phi\ k_1' + \cos\theta\sin\phi\ k_2' - \sin\theta\ k_3')],\\ \frac{\mathrm{d}\phi}{\mathrm{d}t} &= P\cos 2\phi - Q + \frac{R}{\sin\theta}[(-\sin\phi\ k_1 + \cos\phi\ k_2)\cos(\omega t)\\ &\quad + (-\sin\phi\ k_1' + \cos\phi\ k_2')]\\ &\quad |\theta|,\quad |\phi| \le \pi/2. \end{aligned} \tag{3}$$

It is noticed that the modified system of equations reduces to (2) when $k_1' = k_2' = k_3'$ = 0.

The application of an additional external forcing (constant) to control a chaotic system volving under the effect of external forcing (periodic) has been reported in the literature Lakshmanan & Murali 1996). In this method the constant force is applied continuously. Our control strategy involves the application of the constant force for a period of finite ength, T after a period of length, $(n-1)T$ (where $T = 2\pi/\omega$ is the fundamental period f the periodic forcing). Then, we lift the control for a period of finite length, $(n-1)T$ nd thereafter we apply the constant force of the same magnitude for a period of length, ' and again we lift the control for a period of finite length, $(n-1)T$. This process is epeated. Under this process, the system is allowed to evolve according to the system of 2) upto the $(n-1)$th forcing period and evolve according to the system of (3) at the nth orcing period. This process is repeated every nth period. While solving (3), this idea can e implemented by setting $k_1' = k_2' = k_3' = 0$ if $j \not\equiv 0 \bmod(n)$ and $k_1' = k_2' = k_3' \neq 0$ $j \equiv 0 \bmod(n)$, where j represents iteration number and $T = 2\pi/\omega$ is the time interval etween two successive iterations. We note that the above equations for $\dot{\theta}$ and $\dot{\phi}$ decouple n the absence of k_1, k_1' and k_2, k_2'. Hence the presence of an external force field with either , or k_2 is necessary to obtain chaotic solutions in this system. In our calculations we kept

stabilized to a periodic orbit of period equal to an integral multiple of n. On the other hand, if the control is applied continuously, we lose the flexibility of controlling the system to an orbit of desired period.

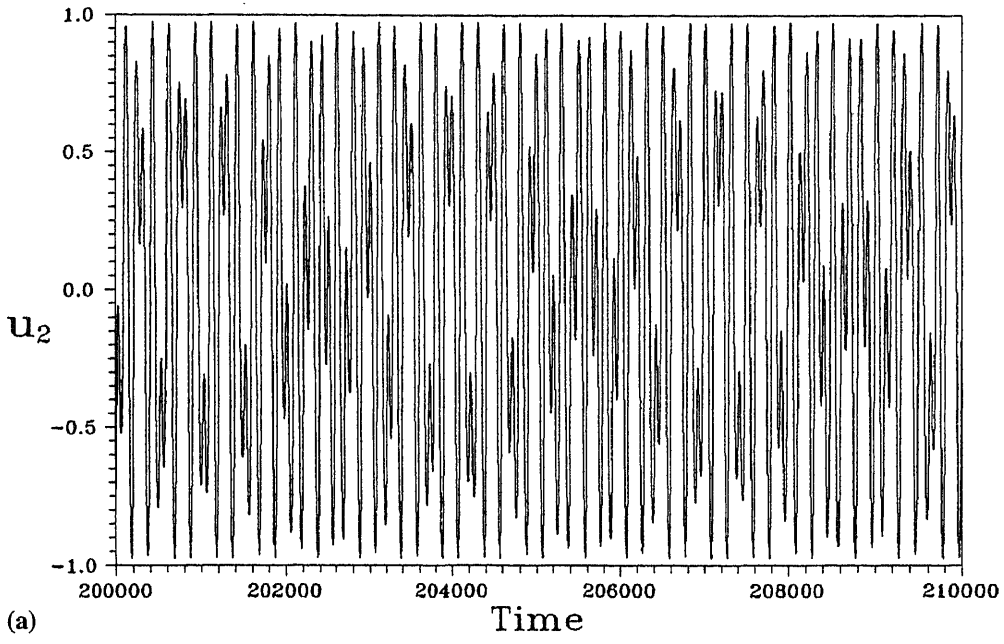
3. Results and discussion

Recently Kumar *et al* (1995) have analysed the properties of the system without control. They tentatively identified chaotic regimes of the parameter k_2 keeping $k_1 = k_3 = 0$. As a first step in analysing the properties of the equations derived in this paper, we set $k_1 = k_3 = 0$, $k_1' = k_2' = k_3' = 0$ and varied k_2 for particles of different aspect ratio within the range of r_e ranging from 0.2 to 2.0 in steps of 0.2 and kept ω equal to J. We expect the greatest complexity of the solutions of these equations for this particular choice of parameters, since k_2 is responsible for the greatest opposition to the hydrodynamic torque due to the imposed shear flow field. At $k_2 = 0$, Jeffery's (1922) results are reproduced and all solutions of the equations starting from different initial conditions tend towards a fixed point in the stroboscopic plot.

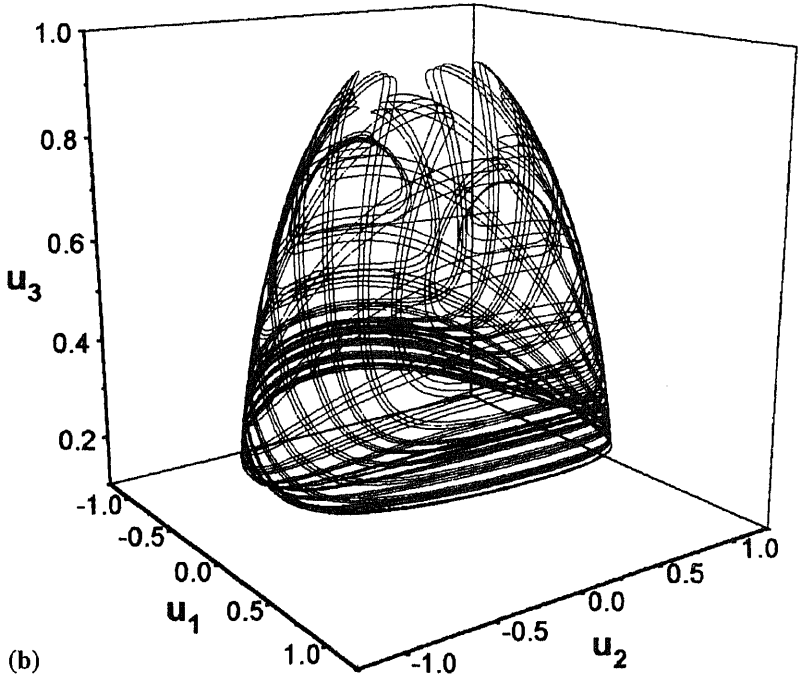
We selected $k_2 = 12.0$ and computed the solution of the equations for different r_e ranging within 0.2 to 2.0 in steps of 0.2. For $k_2 = 12.0$ the system given by (2) behaved chaotically for all the r_e considered except for r_e equal to 1.8 and 2.0. The Lyapunov exponents of the attractors were evaluated and found to be positive. The time series and attractor of a typical trajectory are shown in figures 2a and b for the case of $r_e = 1.2$.

It was reported by Kumar *et al* (1995) that the results of their computations are very sensitive to the aspect ratio of the particle in some parameter regimes. In the case of constant external fields in the same parametric regimes they obtained regular behaviour. For $r_e > 1.0$, they obtained nearly the same fixed point for all initial conditions in the case of a constant force field. This indicates that in the sample application considered in this work, a periodic force field is necessary to effect particle separation for particles with $r_e > 1.0$. Tables 1, 2 and 3 give a sample of the results obtained for a periodic force of amplitude $k_2 = 12.0$, a constant force of amplitude $k_2 = 12.0$ and for zero force respectively. In this case independent separation of particles is possible only for particles of aspect ratio r_e equal to 1.2. The existence of chaotic dynamics in this system allows control of its dynamics to a desired orbit and thus suggests the possibility of better separation of particles for almost all particles of aspect ratio ranging within 0.2 to 2.0. This is difficult to obtain in the case of regular behaviour or chaotic behaviour without control.

We stabilized the chaotic dynamics obtained at $k_2 = 12.0$ to a periodic behaviour of period equal to a multiple of the fundamental period, $2\pi/\omega$ to demonstrate the applicability of the method proposed. The time series of u_2 of the Poincaré section (stroboscopic plot) computed from (2) for $r_e = 1.6$ is given in figure 3. It is demonstrated that the system can be controlled to periodic behaviour of any desired period by applying the same constant force. For example, the system could be controlled to period-2, period-3, period-4 and period-5 orbits by applying the same constant force equal to $k_2' = 5$ as can be seen from figures 4, 5, 6 and 7. This is obtained by setting $k_1' = k_2' = k_3' = 0$ if $j \not\equiv 0 \bmod(n)$ and $k_1' = k_3' = 0$, $k_2' = 5.0$ if $j \equiv 0 \bmod(n)$ in the evolution of (3) when the system is

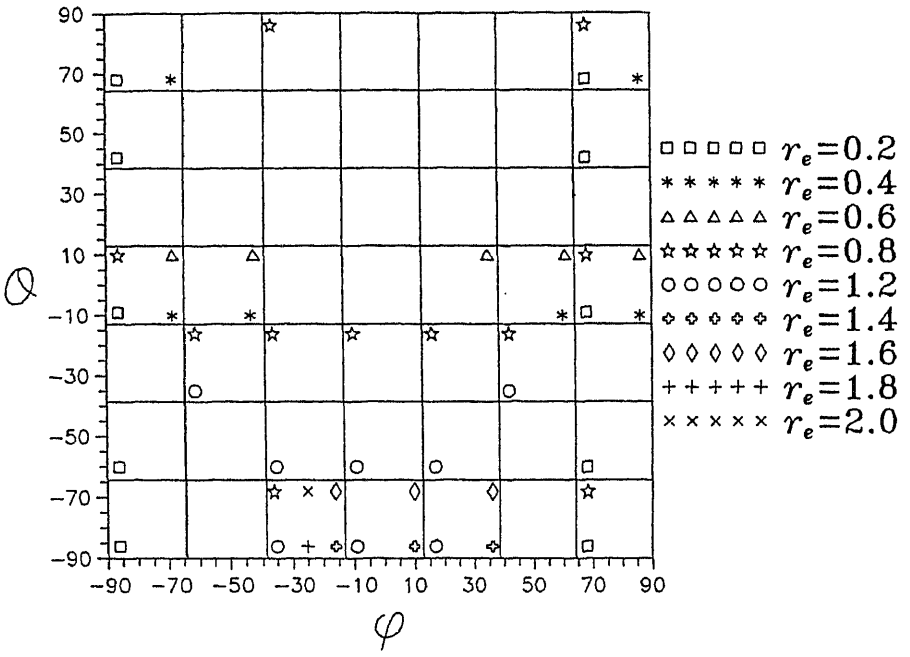

1.0
0.5
0.0
−0.5
−1.0
u_2
200000
202000
204000
206000
208000
210000
Time

(a)


1.0
0.8
0.6
0.4
0.2
u_3
-1.0
-0.5
0.0
0.5
1.0
u_1
1.0
0.5
0.0
-0.5
-1.0
u_2

(b)

Table 1. Distribution of evolution of initially uniformly distributed particles of different aspect ratios for the case $k_2 = 12$; $\omega = J$; $5 \leq m \leq 49$ and $100 \leq l \leq 1000$, where m is the total number of particles in the grid on the average and l is the total number of occurrences of the grid.

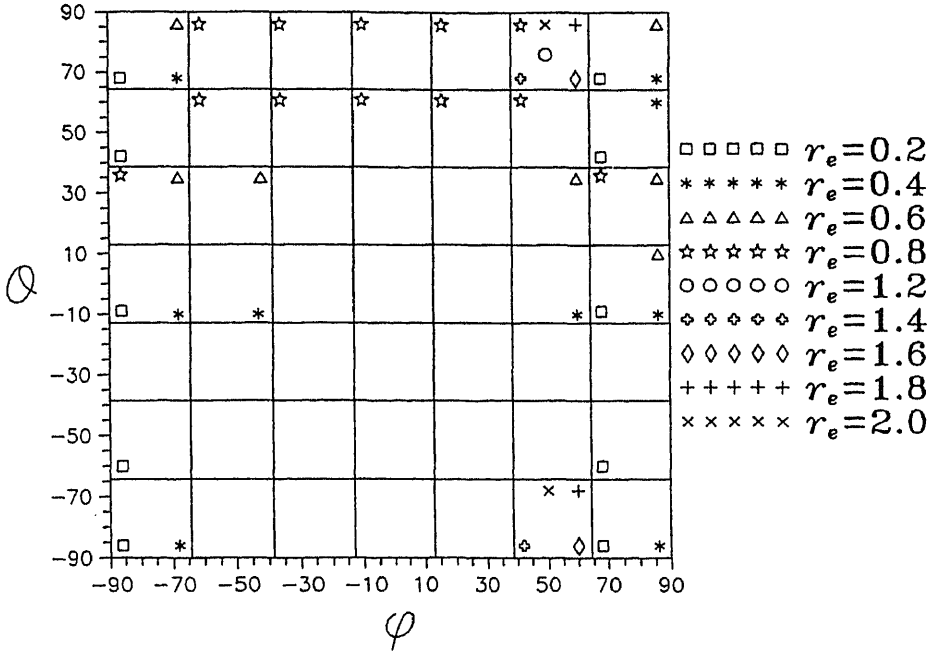


to be stabilized in a period-n orbit. The system evolves according to the evolution (2) for $j \not\equiv 0 \bmod(n)$ and according to the evolution (3) for $j \equiv 0 \bmod(n)$. It is also possible to stabilize a given chaotic system to different n-period orbits by suitably changing the control parameter k_2'. As an example, the time series of Poincaré sections corresponding to two different 2-period solutions embedded in a chaotic system are given in figure 8.

The control technique is applied after 2000 fundamental periods and the system is followed upto 7000 periods with control. Once control is applied the system stabilizes rapidly to the appropriate periodic orbit. The magnitude of the perturbation required to stabilize the system was small compared to the magnitude of the periodic force for all aspect ratios except $r_e = 0.2$, 0.4 and 0.6. The magnitude of the constant force required for control depends on the aspect ratio of the particle and the desired period, n for aspect ratios less than or equal to 0.6.

One important advantage of the control algorithm outlined in this work is the possibility of switching over to different periodic solutions during a given run. This implies that a system in chaotic dynamics can be stabilized to one particular periodic orbit for a given time and to another periodic orbit of entirely different period after a given time as shown in figure 9. In this example the control parameter $k_2' = 5.0$ is applied between 2001 and 3000 periods to stabilize the system to a period-2 orbit. Then the control is lifted between 3001 and 4000 periods and again the control parameter $k_2' = 5.0$ is applied between 4001 and

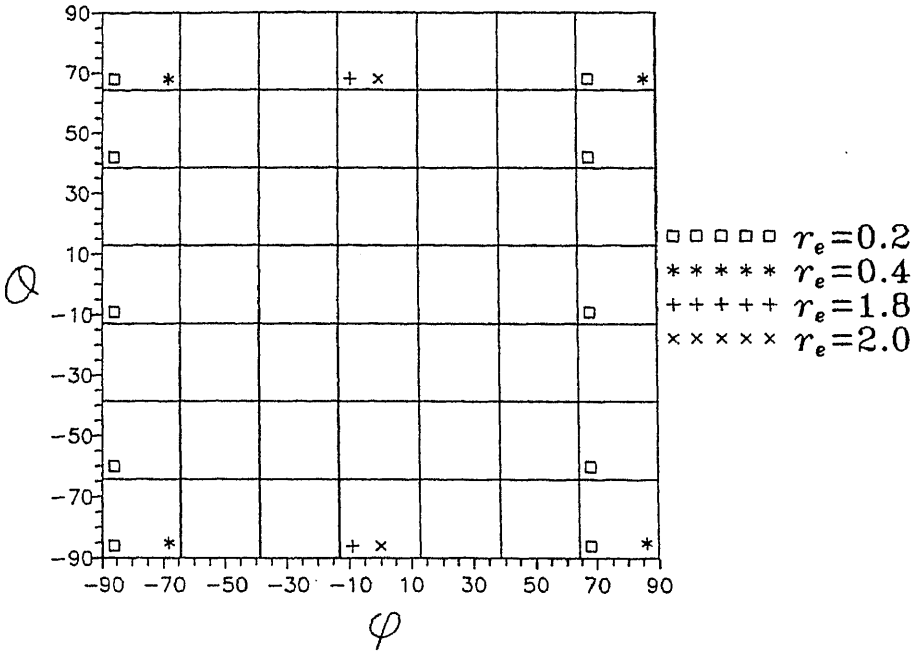
in the grid on the average and l is the total number of occurrences of the grid.



5000 periods to stabilize the system to a period-3 orbit. The control was removed again between 5001 and 6000 periods. Thus the system oscillates in a period-2 orbit between 2001–3000 periods and then oscillates in a period-3 orbit between 4001–5000 periods as can be seen from figure 9. This figure also reveals the fact that once the control is lifted, the system returns to a chaotic system. Another advantage which is important from the point of view of the sample application proposed in this work is the possibility of changing the periodic behaviour to another orbit of the same period by suitably changing the magnitude of the applied constant force. This allows us to bring a particle having a definite aspect ratio to a desired orbit by changing control parameters. It is also noted that all initially uniformly distributed particles of a given aspect ratio within the range $0.8 \leq r_e \leq 2.0$ can be concentrated in a given grid by applying a periodic force with control as can be seen from tables 4, 5 and 6.

The observations made in this work suggest the possibility of separating particles more efficiently based on control of chaotic dynamics. We applied a control force (constant force) of magnitude $1 \leq k_2' \leq 5$. For this range of k_2', the chaotic dynamics of particles of all aspect ratio except $r_e = 0.2$, 0.4 and 0.6 could be controlled to a desired orbit. This range of control force seems to be sufficient for efficient separation of particles by shape since particles of these aspect ratios alone can be brought to a desired position. To develop quantitative results based on this observation, we divided the range of possible orientations

Table 3. Distribution of evolution of initially uniformly distributed particles of different aspect ratios for the case $k_2 = 0$; $\omega = 0$; $5 \leq m \leq 49$ and $100 \leq l \leq 1000$, where m is the total number of particles in the grid on the average and l is the total number of occurrences of the grid.



namely $[-90°, 90°]$ in both θ and ϕ variables into 7 equal intervals resulting in 49 equal sized grids. We then computed the evolution of initially uniformly distributed particles of different aspect ratio within the range of r_e equal to 0.2 to 2 in steps of 0.2.

We followed the evolution of the ensemble of particles for 3000 iterations of the stroboscopic plot and deleted the first 1000 values to remove transients in the case of constant and zero forces and in the case of periodic forces with control and without control. In all cases we calculated the number of particles in each grid on every second iteration of the Poincaré section of the evolution equations resulting in a total of 1000 values. We noted the grids in which the total number of particles was greater than or equal to 5 and also noted the number of particles in each grid only if the particle occurred in that grid in more than 100 iterations in all cases. We denote these values as r_e, l, m where r_e, l, m denote the aspect ratio, total number of occurrences of the grid and total number of particles in the grid on the average respectively and prepared tables for these values. A particle of a given aspect ratio visiting a given grid and the absence of particles of other aspect ratios visiting the same grid is more important than the number of visitations of a given grid and the number of particles visiting a grid from the point of view of particle separation. Since our earlier computations indicated the greatest sensitivity of the results to the aspect ratio near $k_2 = 10$, we selected k_2 equal to 12.0 for comparing the effects of periodic forces with and without control and as a magnitude of the constant force.

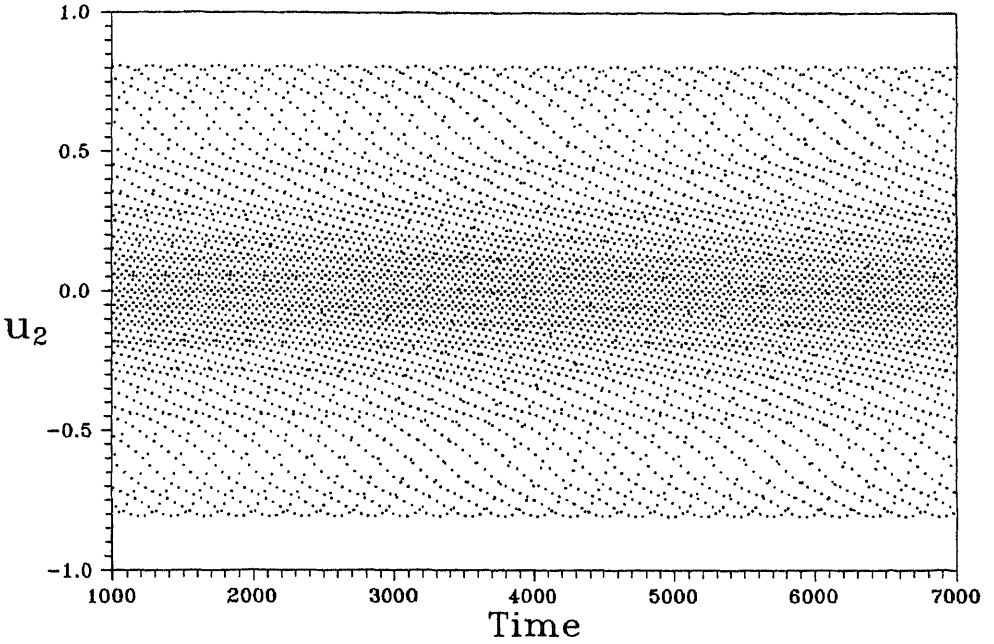


Figure 3. Trajectory plot of u_2 vs time at every intersection of the trajectory with the Poincaré plane for $k_2 = 12$; initial conditions $\theta = \phi = 45°$; $\omega = J$; $r_e = 1.6$.

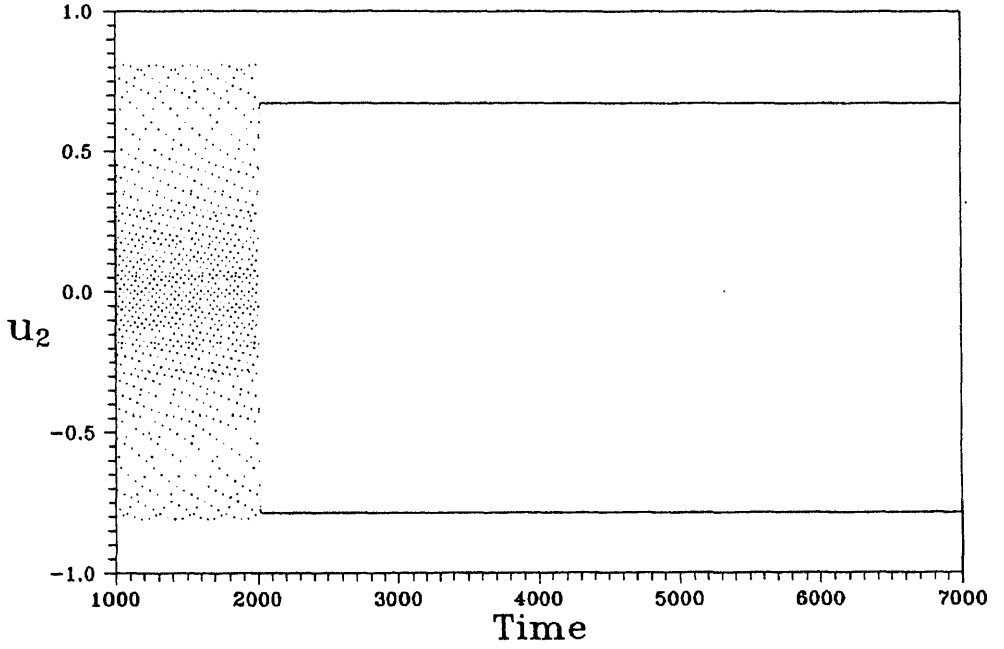


Figure 4. Stroboscopic plot with control showing a period-2 solution for $k_2 = 12$; $k_2' = 5.0$; initial conditions $\theta = \phi = 45°$; $\omega = J$; $r_e = 1.6$.

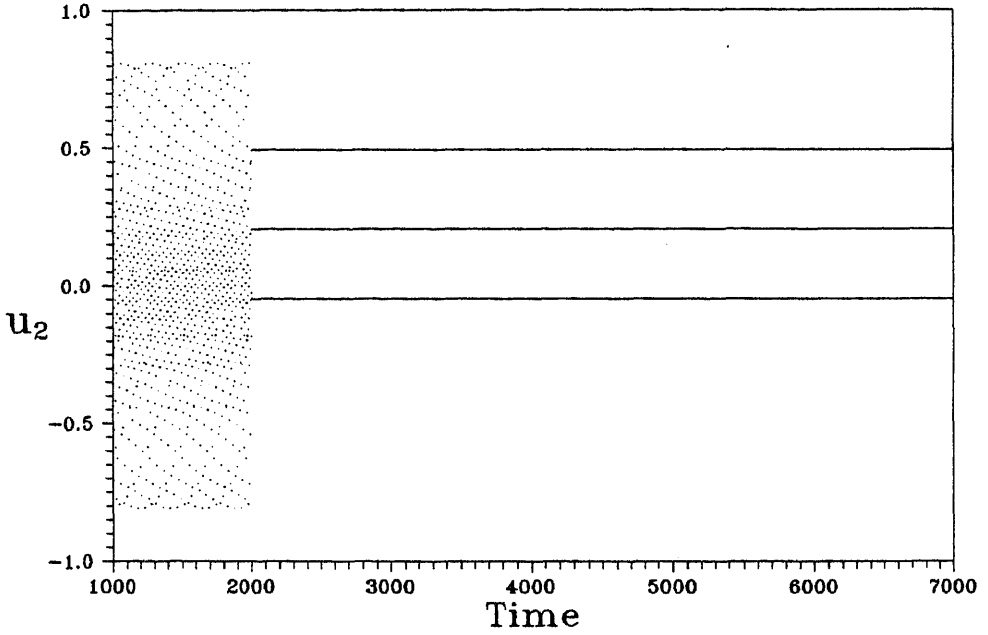


Figure 5. Stroboscopic plot with control showing a period-3 solution for $k_2 = 12$; $k_2' = 5.0$; initial conditions $\theta = \phi = 45°$; $\omega = J$; $r_e = 1.6$.

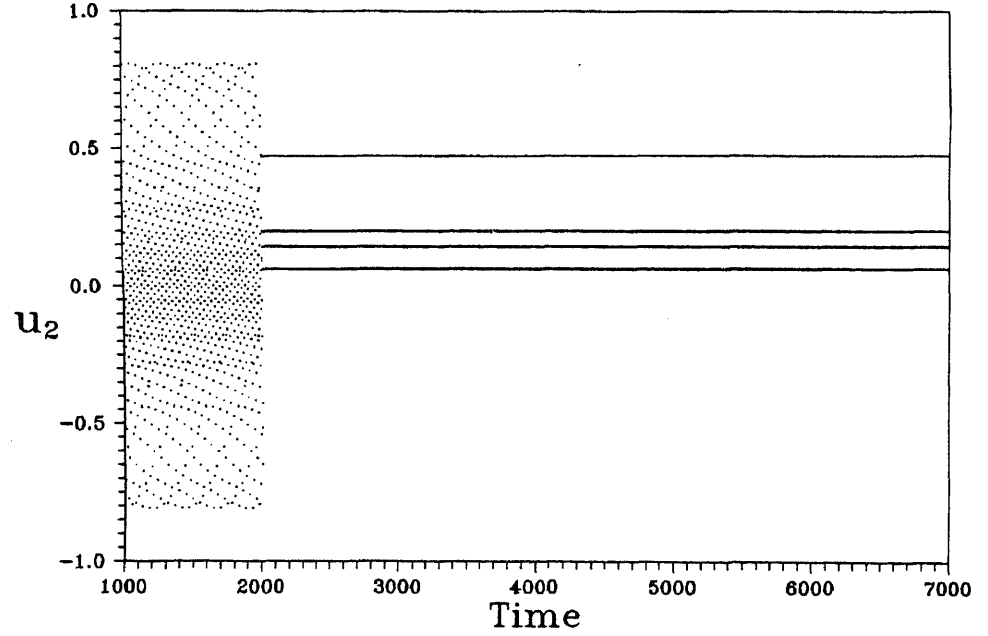


Figure 6. Stroboscopic plot with control showing a period-4 solution for $k_2 = 12$; $k_2' = 5.0$; initial conditions $\theta = \phi = 45°$; $\omega = J$; $r_e = 1.6$.

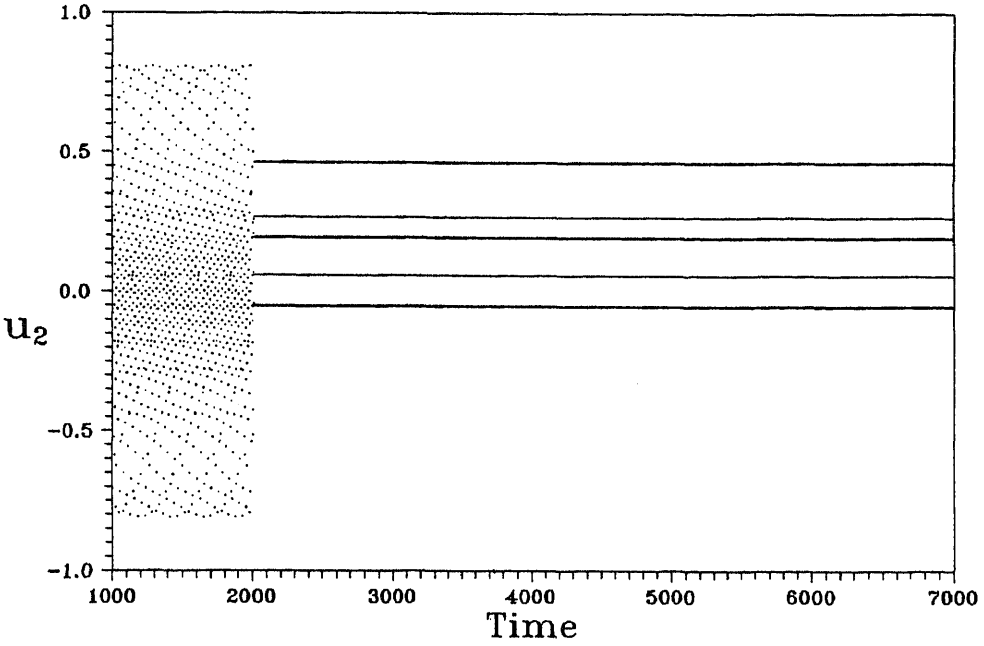


Figure 7. Stroboscopic plot with control showing a period-5 solution for $k_2 = 12; k_2' = 5.0$; initial conditions $\theta = \phi = 45°; \omega = J; r_e = 1.6$.

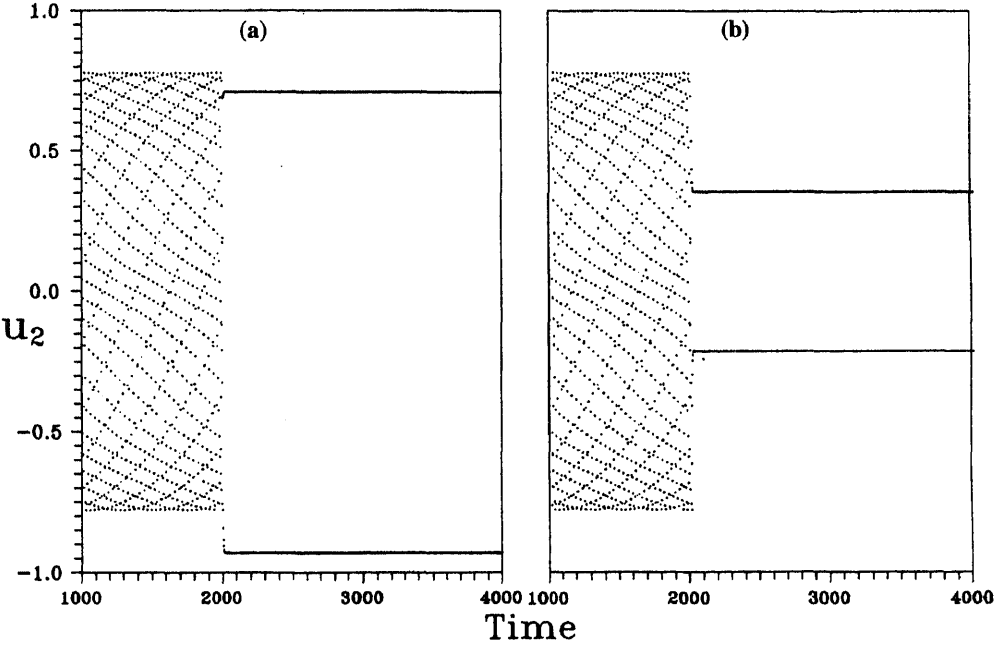


Figure 8. Stroboscopic plot with different controls showing two different period-2 orbits obtained for $k_2 = 12$; initial conditions $\theta = \phi = 45°; \omega = J; r_e = 1.2$ with control **(a)** $k_2' = 4.0$, **(b)** $k_2' = 5.0$.

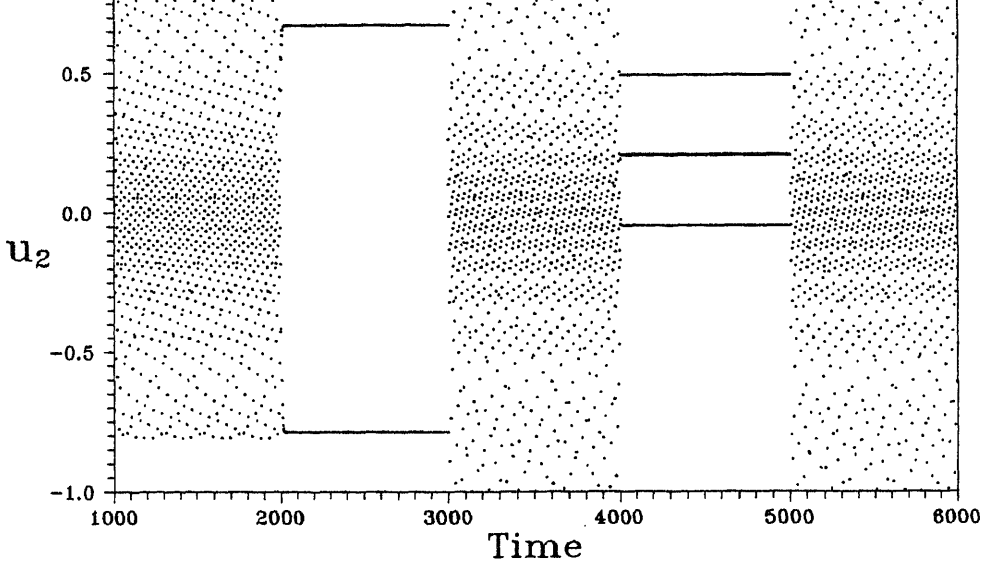
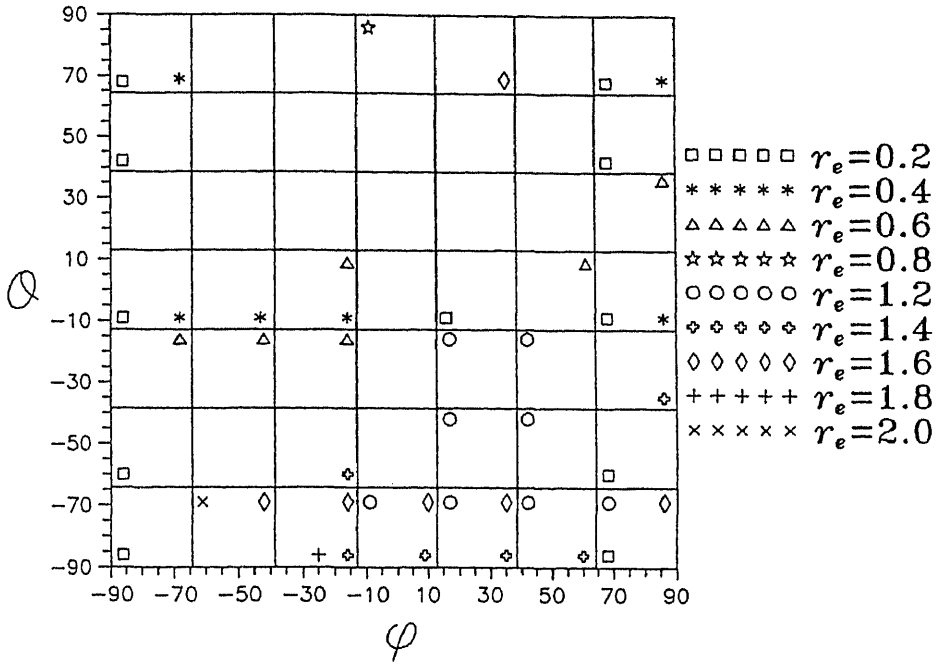


Figure 9. Stroboscopic plot showing period-2 and period-3 orbits successively obtained with control applied at every second and third periods respectively for $k_2 = 12$; initial conditions $\theta = \phi = 45°$; $\omega = J$; $r_e = 1.6$; $k_2' = 5.0$.

We noted the evolution of the ensemble of particles at every 2nd iteration. We analysed in detail the cases of periodic force with and without control and compared the results with the constant force and zero forces. The fact that a chaotic behaviour can be controlled to oscillate in a selected period resulted in more efficient separation enabling all initially and uniformly distributed particles to be directed to a desired grid by a suitably engineered control technique. For example, all the 49 initially and uniformly distributed particles of aspect ratio $r_e = 1.2$ could be brought to the grids $(1, 4)$ or $(7, 4)$ by applying $k_2' = 4$. Under the same circumstances it was possible to bring all the 49 particles of aspect ratio $r_e = 1.2$ to the grids $(1, 3)$ or $(7, 1)$ by applying a control of constant force $k_2' = 5$. A sample of the results we obtained for a periodic force of magnitude $k_2 = 12$ with control forces $k_2' = 2$, 3 and 5 and for constant and periodic forces of magnitude $k_2 = 12.0$ and zero forces are given in tables 1–6. The tables presented in this paper indicate that controlling chaotic dynamics is preferable to chaotic behaviour and regular behaviour for efficient separation of particles.

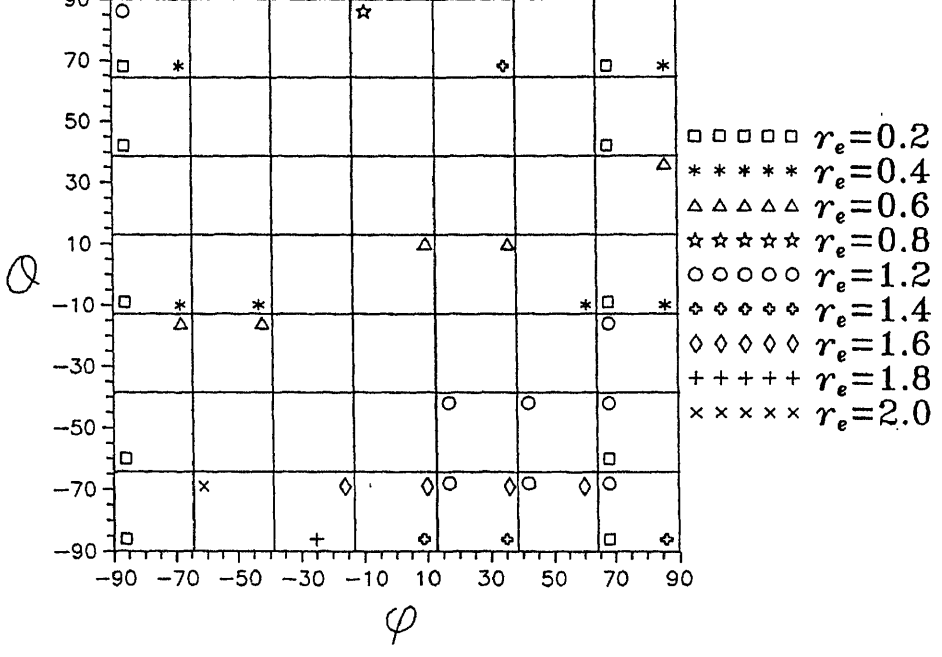
A detailed analysis of all the tables indicates that particles of aspect ratio 0.2 alone can be separated from a mixture containing particles of different aspect ratios ranging within 0.2 to 2.0 by applying a control force of magnitude $k_2' = 1, 2, 3, 4$ or 5 applied at every second period along with $k_2 = 12$, since particles of aspect ratio 0.2 alone occur in some grids as can be seen from the sample tables. In the case of periodic forces without control and constant and zero forces particles of aspect ratio 0.2 alone also occur along the boundary of the tables. Kumar *et al* (1995) suggested that particles of this aspect ratio alone could be separated by applying a constant force. However, in the case of a periodic force with

Table 4. Distribution of evolution of initially uniformly distributed particles of different aspect ratios for the case $k_2 = 12$; $k_2' = 2$; $\omega = J$; $5 \leq m \leq 49$ and $100 \leq l \leq 1000$, where m is the total number of particles in the grid on the average and l is the total number of occurrences of the grid.



control, the occurrences of this particle alone are concentrated among a fewer number of grids and visit a given grid a larger number of times. In the case of zero forces the particles of this aspect ratio alone are concentrated among a fewer number of grids along with particles of higher aspect ratio. Thus, for the separation of the particles of this aspect ratio alone a periodic force of magnitude $k_2 = 12.0$ with control is preferable to any other possibility. Similar analysis of the tables shows that it is desirable to apply a periodic force with control for separating particles of aspect ratio 0.4. As can be seen from the sample tables particles of aspect ratio 0.6 alone and 0.8 alone also can separated individually by applying a periodic force with control. In the case of particles of aspect ratio 0.8, they can be brought to one of the extreme grids. Thus, a periodic force with control seems to be preferable to any other case considered for effective separation of particles in the case of aspect ratios $r_e < 1$.

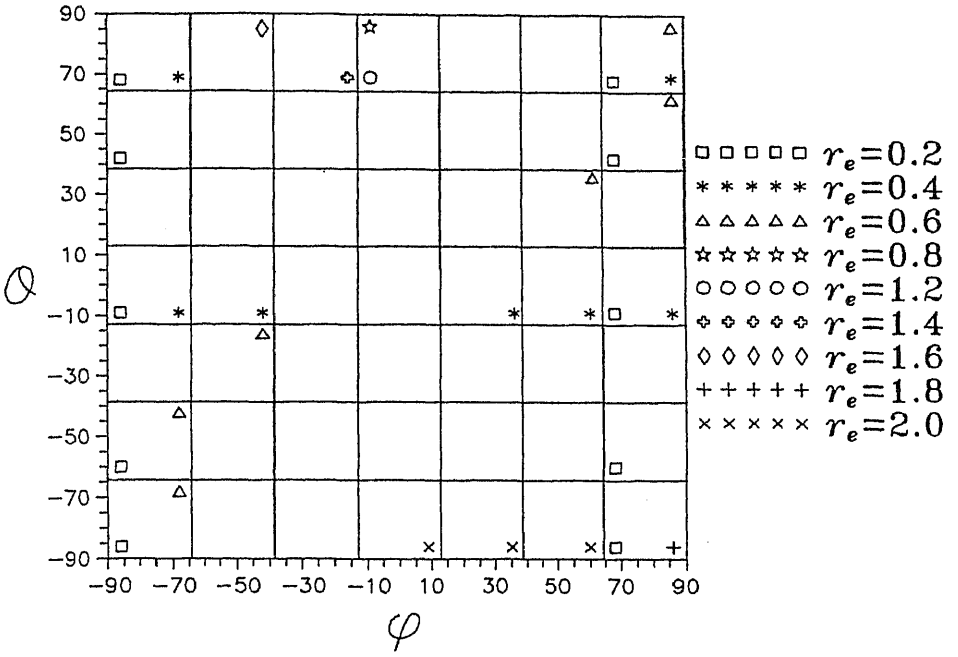
Kumar *et al* (1995) demonstrated that periodic forces are preferable to constant forces and zero forces for separating particles of aspect ratio $r_e > 1$. In their analysis they observed that the occurrence of particles of aspect ratio $r_e > 1$ are spread among a large number of grids. However, for particles of aspect ratio greater than or equal to 1.6 individual separation was not possible in their analysis. Further study of the tables prepared in this work indicate that a periodic force with control is once again preferable in such cases. In the case of particles of aspect ratio $r_e > 1.0$ all the initially uniformly distributed particles of the same aspect ratio could be brought to one grid except for $r_e = 1.8$. Hence particles of



aspect ratio within the range of r_e ranging from 1.2 to 2.0 except for $r_e = 1.8$ can be easily separated as can be seen from the tables. For particles of aspect ratio equal to 1.8 individual separation may not be possible in the cases considered in this work since particles of this aspect ratio appear in combination with particles of lower aspect ratio equal to 0.2 in some cases considered with control. Hence if particles of aspect ratio 0.2 have been separated out as explained earlier, particles of aspect ratio equal to 1.8 can be separated from the mixture. One advantage of periodic forces with control is that all particles of aspect ratio 1.8 can be brought to one grid in combination with the particles of aspect ratio 0.2 as shown in table 6.

In conclusion, it has been generally noted that control of chaotic behaviour gives better separation than chaotic and regular behaviour. One of the main features of the method suggested in this paper is that all particles of the same aspect ratio can be concentrated in a previously desired grid. A detailed analysis of the problem can suggest suitable designs for separation of particles by aspect ratio to get a well-characterized suspension of particles. A possible design for this separation of particles from a mixture of particles of different aspect ratios based on the differences in the orientation of the particle may consist of a base plate having grooves along different orientations so that when the particles are oriented in such directions, they settle in a particular groove and can be separated out at every integral multiple of the fundamental period.

particles in the grid on the average and I is the total number of occurrences of the grid.



Thus, even in the rather simple application considered in this work, control of chaos leads to greater efficiency. This suggests that the possibility of chaos control should be important in many of the applications mentioned in the introduction. The novel control of chaos technique suggested in this paper has been demonstrated to be effective even in the relatively complex problem considered here. An additional feature of the control of chaos technique suggested in this paper is that the control is effected very rapidly and the behaviour of the concerned system can be switched from one desired period to another desired period very quickly. This suggests that this control of chaos technique may be applied to other chaotic systems very effectively.

The authors wish to thank Dr A D Damodaran, and Dr B C Pai, Regional Research Laboratory (CSIR), Thiruvananthapuram for their constant encouragement during the work. The authors also wish to thank Dr S Savithri, Mr C S Bhat and Mr K Satheesh Kumar for useful discussions and suggestions. One of the authors (C V Anil Kumar) wishes to acknowledge the Council of Scientific and Industrial Research, India for financial support. Financial support from the Department of Science and Technology, Government of India, Grant No. SP/S2/R-04/93 is also gratefully acknowledged.

b	equatorial radius;	[L]
$\mathbf{i}$	unit vector in the x direction;	[–]
J	$2\pi(r_e + {r_e}^{-1})$	[–]
$\mathbf{k}$	periodic force field;	$[MLT^{-2}]$
$\mathbf{k}'$	constant force field;	$[MLT^{-2}]$
${}^c\hat{\mathbf{K}}$	intrinsic properties of the particle, which depend upon the geometric configuration of its wetted surface;	$[L^2]$
${}^r\hat{\mathbf{K}}$, $\hat{\tau}$	intrinsic properties of the particle, which depend upon the geometric configuration of its wetted surface;	$[L^3]$
l	total number of occurrences of the grid;	[–]
$\mathbf{L}$	torque about the centre of the particle exerted by the fluid on the particle;	$[ML^2T^{-2}]$
m	total number of particles in the grid on the average;	[–]
n	period of controlling;	[–]
r_e	particle aspect ratio given by $r_e = a/b$;	[–]
$\mathbf{S}$	rate of deformation tensor given by $\mathbf{S} = \frac{1}{2}[\nabla\mathbf{V} + \nabla\mathbf{V}^T]$;	$[T^{-1}]$
$\mathbf{u}^*$	unit vector along the axis of symmetry which determines the orientation of the spheroids with components $\mathbf{u}_1{}^*$, $\mathbf{u}_2{}^*$, $\mathbf{u}_3{}^*$;	[–]
$\mathbf{U}$	translational velocity;	$[LT^{-1}]$
$\mathbf{U} - \mathbf{V}$	translational slip velocity;	$[LT^{-1}]$
y	y coordinate;	[L]
$\dot{\gamma}$	shear rate;	$[T^{-1}]$
μ_0	viscosity of the fluid;	$[ML^{-1}T^{-1}]$
θ, ϕ	polar angle and the azimuthal angle corresponding to a given vector $\mathbf{u}$ respectively;	[–]
$\dot{\theta}$	time derivative of θ;	[–]
$\dot{\phi}$	time derivative of ϕ;	[–]
ω	frequency of the driving force;	$[T^{-1}]$
$\mathbf{\Omega}$	angular velocity of the particle;	$[T^{-1}]$
$\mathbf{\Omega} - \boldsymbol{\omega}$	rotational slip velocity;	$[T^{-1}]$
∇	del operator.	$[L^{-1}]$

References

Buevich Yu A, Syutkin S V, Tetyukhin V V 1984 Theory of a developed magnetofluidized bed. *Magnetohydrodynamics* 20: 333–339

Cebers A 1993a Chaos in polarization relaxation of a low-conducting suspension of anisotropic particles. *J. Magn. Magn. Mater.* 122: 277–280

Cebers A 1993b Chaos: new trend of magnetic fluid research. *J. Magn. Magn. Mater.* 122: 281–285

Christini D J, Collins J J 1995 Controlling nonchaotic neuronal noise using chaos control techniques. *Phys. Rev. Lett.* 75: 2782–2785

Ditto W L, Spano M L, Lindner J F 1995 Techniques for the control of chaos. *Physica D* 86: 198–211

Fronzoni L, Giocondo M, Pettini M 1991 Experimental evidence of suppression of chaos by resonant parametric perturbations. *Phys. Rev.* A43: 6483–6487

Güémez J, Gutiérrez J M, Iglesias A, Matias 1994 Stabilisation of periodic and quasiperiodic motion in chaotic systems through changes in the system variable. *Phys. Lett.* A190: 429–433

Ignatenko N M, Melik-Gaikazyan I Yu, Polunin V M, Tsebers A O 1984 Excitation of ultrasonic vibrations in a suspension of uniaxial ferromagnetic particles by volume magnetostriction. *Magnetohydrodynamics* 20: 237–240

Jeffery G B 1922 The motion of ellipsoidal particles immersed in a viscous fluid. *Proc. R. Soc.* A102: 161–199

Kashevskii B É 1986 Torque and rotational hysteresis in a suspension of single-domain ferromagnetic particles. *Magnetohydrodynamics* 22: 161–168

Kumar C V A, Kumar K S, Ramamohan T R 1995 Chaotic dynamics of periodically forced spheroids in simple shear flow with potential application to particle separation. *Rheol. Acta* 34: 504–512

Lakshmanan M, Murali K 1996 *Chaos in nonlinear oscillators: Controlling and synchronization* (Singapore: World Scientific)

Ott E, Grebogi C, Yorke J A 1990 Controlling chaos. *Phys. Rev. Lett.* 64: 1196–1199

Petrikevich A V, Raikher Yu L 1984 Rheological characteristics of magnetic suspension in alternating magnetic field. *Magnetohydrodynamics* 20: 122–127

Rajasekar S, Lakshmanan M 1992 Controlling of chaos in Bonhoeffer-van der Pol oscillator. *Int. J. Bifurcation Chaos* 2: 201–204

Rajasekar S, Lakshmanan M 1993 Algorithms for controlling chaotic motion: application for the BVP oscillator. *Physica D* 67: 282–300

Rhode M A, Rollins R W, Markworth A J, Edwards K D, Nguyen K, Daw C S, Thomas J F 1995 Controlling chaos in a model of thermal pulse combustion. *J. Appl. Phys.* 78: 2224–2232

Shul´man Z P, Khusid B M, Zal´tsgendler E A 1986 Motion of an axisymmetric magnetically soft particle in hydrodynamic flow under the action of a strong rotating magnetic field. *Magnetohydrodynamics* 22: 288–293

Tsebers A O 1986 Numerical modelling of the dynamics of a drop of magnetizable liquid in constant and rotating magnetic fields. *Magnetohydrodynamics* 22: 345–351

Zibol´d A F, Kapusta A B, Keskyula V F, Petrov G N, Remizov O A 1986 Hydrodynamic phenomena accompanying the growth of single crystals by Czochralski's method in a rotating magnetic field. *Magnetohydrodynamics* 22: 202-206

Zumbrunnen D A, Miles K C, Liu Y H 1996 Auto-processing of very fine-scale composite materials by chaotic mixing of metals. *Composites* A27: 37–47

Flexible regenerative supervision of sequential behaviour

A K RAINA [1] and A G VALAVI [2]

[1] Department of Electrical Engineering, Indian Institute of Technology, Kanpur 208 016, India
[2] Wipro Limited (Ist floor), 37 Castle Street, Ashoknagar, Bangalore 560 025, India
e-mail: akraina@iitk.ernet.in; valavi@wipinfo.soft.net

MS received 14 June 1996; revised 14 May 1997

Abstract. This paper investigates the non-terminating (NTB) behaviour of discrete-event processes (DEPs) modelled by finite state Büchi automata. Specifically, we develop synthesis criteria and procedures for state-feedback supervisors that ensure infinitely often (i.o.) returns to a given *good* state subset of the DEP in the closed loop. This problem is referred to as the *stabilization* problem for discrete event systems (DESs) elsewhere.

The approach of the present paper is distinguished by the fact that the stabilization is achieved without major compromises on the richness of behaviour, or manouverability. This flexibility comes as a consequence of using two supervisors in parallel and switching them on-line to achieve the desired behaviour characteristics in a non-blocking manner. Supervision is arrived at by decomposing the state space into *controllable* and *uncontrollable* subspaces relative to the desired target space, much in a manner identical to control theory for systems modelled by differential equations.

Keywords. Regenerative supervision; non-terminating behaviour; discrete event processes.

1. Introduction

Supervisory control of discrete-event systems (DESs) modelled as finite-state machines (FSMs) is an important control research area presently. The basic model and control theory, in an FSM-Formal Language set-up, are due to Ramadge and Wonham (Ramadge & Wonham 1987; Wonham & Ramadge 1987). Initially, the supervisory theory was developed only for the transient behaviour control of a DES, which behaviour arises, naturally, from simple FSM models and the *regular* finite-string behaviour sets of such models (Eilenberg 1974; Hopcroft & Ullman 1979). Later the theory was extended by Ramadge (1989) and others (Kumar *et al* 1992) to control the steady-state DES behaviour by using Büchii

extensions of FSMs, or the so-called Büchi automata (Eilenberg 1974), which model non-terminating (or sequential) response of the FSMs.

An important problem that arises in the study of sequential DES behaviour is that of stability. Özveren & Willsky (1991) define *stability* as infinitely often (i.o.) guaranteed returns, by means of a state feedback, to a set of *a priori* selected *good* states of the DES, in a non-Büchii setting. This paper addresses a similar problem in a Büchii setting which, as mentioned above, arises as the natural extension of FSMs (Eilenberg 1974) to model sequential behaviour. Furthermore, the control strategies developed here ensure *flexibility* in the sense that the rate of convergence to the *good* set can be varied on-line, by adjusting a single free parameter, so that depending upon other (possibly unmodelled) conditions one can trade-off richness of behaviour (manouverability) for quick stabilization. This is done by having two logically connected parallel controllers which are switched based on the requirement at any given instant during the operation or evolution of the process. Moreover, the controllers we develop satisfy the required minimal-*restrictiveness* and non-*blocking* conditions (Ramadge & Wonham 1987), so that the closed loop behaviour is maximally-*permissible* (Wonham & Ramadge 1987) and satisfies *liveness* conditions (Özveren & Willsky 1987). Another important feature is that finite and sequential – that is transient and steady-state – specifications are simultaneously handled, thus simplifying the overall control process. In addition, the stringent requirement of topological closure of the specification behaviour (Ramadge 1989) is avoided here.

There are other notions of *stability* as well. For example, Brave & Heyman (1989) consider the *good* set as an *attractive* invariant domain in the closed loop under state feedback. The control problem then is to design a feedback so that the DES enters the given set in a finite number of steps and does not exit it later. This could, obviously, constrain the closed loop behaviour seriously since it compromises the *richness* of behaviour. Other notions of *stability* that involve the sublanguages of the behaviour of a DES (Kumar *et al*) rather than state subsets of the underlying FSM also exist. We shall, however, deal only with the state-set defined case, and call our *flexible* stabilizers *regenerative* supervisors, to highlight the fact that the controller ensures, in the closed loop, i.o. visitations to the *a priori* chosen *good* state subset.

This paper is organised as follows: § 2 develops the necessary formal background for DEPs and their supervisors in terms of both state dynamics and language behaviour. The main section of this paper is § 3, which lays down the synthesis theory for design of the required flexible supervisors. An important part of this section is the proof of the necessity condition of theorem 1, which is separately given as appendix A due to importance beyond the needs of the theorem. Section 4 concludes the paper by highlighting the advantages of our procedure with the help of a near generic example given in appendix B.

2. Mathematical preliminaries

2.1 *Languages*

Let Σ be a finite set and Σ^* the set of all finite strings over Σ. Σ^* is a free *monoid* (Eilenberg 1974) under *catenation* with the empty string ϵ being the monoid identity. For $s \in \Sigma^*$, $|s|$

denotes the length of the string s, so that $|\epsilon| = 0$. For $u, v \in \Sigma^*$, we say that u is a *prefix* (proper *prefix*) of v, denoted by $u \preceq v$ ($u \prec v$), if for some $\omega \in \Sigma^*$, $v = u\omega$ ($|\omega| > 0$). Clearly the relation '$\preceq$' is a partial order on Σ^*.

A subset $L \subseteq \Sigma^*$ is called a *Language* over Σ, and the set $\bar{L} = \{u | uv \in L \text{ for some } v \in \Sigma^*\}$ is called the prefix closure of L (Eilenberg 1974). L itself is prefix closed if $\bar{L} = L$. Next we define ω-Languages (Eilenberg 1974; Ramadge 1989) which constitute infinte or sequential extension of strings.

Let N denote the natural number set and let Σ^ω denote the set of sequences over Σ, i.e.,

$$\Sigma^\omega \doteq \{e | \, e : N \longrightarrow \Sigma\}. \tag{1}$$

For $e \in \Sigma^\omega$ and $j \in N$, let $e(j)$ denotes the jth element of e, and $e^j = e(1)e(2)\ldots e(j)$ denote the finite string (i.e., $e^j \in \Sigma^*$) consisting of first j elements of e. In general, $u \in \Sigma^*$ will be called a prefix of $e \in \Sigma^\omega$ (denoted, once again, by $u \preceq e$) if $u = e^j$ for some $j \in N$. $B \subseteq \Sigma^\omega$ is called an ω-Language over Σ, and the language $pr(B) \in \Sigma^*$ defined as;

$$pr(B) = \{\epsilon\} \cup \{e^j | \, e \in B, j \geq 1\}, \tag{2}$$

is called the *prefix* of B.

For a monotonic sequence $u_1 \prec u_2 \prec u_3 \cdots$ of elements over Σ^*, $\exists$ a unique element $e \in \Sigma^\omega$ such that $e^j = u_k$, for $j = |u_k|, k \in N$ (Eilenberg 1974). Such an e is called the limit of the sequence $\{u_k\}$. The *adherence* or *limit* of a prefix-closed $L \subseteq \Sigma^*$ is the ω-Language L^∞ defined as;

$$L^\infty = \{e | e \in \Sigma^\omega \wedge e^j \in L, \ \forall j \in N\}. \tag{3}$$

If L is not prefix closed it suffices that $e^j \in L$ for infinitely many $j \in N$ (Eilenberg 1974; Ramadge 1989).

A crucial metric topology on Σ^ω is induced by the metric ρ, defined as under (Eilenberg 1974);

$$\rho(e_1, e_2) = 1/n \text{ if } e_1^{n-1} = e_2^{n-1} \text{ and } e_1^n \neq e_2^n; \text{ and,} \tag{4}$$

$$\rho(e_1, e_2) = 0 \text{ if } e_1^n = e_2^n. \tag{5}$$

It is easy to show that in this topology Σ^ω is *compact* (Eilenberg 1974). Furthermore, this metric can be extended in a straightforward manner to $\Sigma^* \cup \Sigma^\omega$, though we shall not need that extension here in the sequel. The topological closure $\bar{B}$ (B union its limit points) of an ω-Language B equals the limit of its prefix language, i.e., $\bar{B} = [pr(B)]^\infty$, as can be readily verified (Eilenberg 1974; Ramadge 1989). For $B \subseteq S \subseteq \Sigma^\omega$, we say B is closed relative to S if $\bar{B} \cap S = B$ (Ramadge 1989).

2.2 *Discrete-event systems*

A discrete-event system (Ramadge & Wonham 1987) is modelled as a finite state automaton (Eilenberg 1974) $G = (Q, \Sigma, \delta, q_0)$ where Q is the (finite) state set, Σ the finite set of events (also called the process *alphabet*), $\delta : Q \times \Sigma \longrightarrow Q$ gives the state transition (*partial*) function, which can be naturally extended to Σ^* by defining

$\delta(q, \epsilon) = q$, and $\delta(q, s\sigma) = \delta(\delta(q, s), \sigma)\ \forall s \in \Sigma^*$, whenever all intermediate transitions are defined; else it will be undefined. Whenever $\delta(q, s)$ is defined for some $s \in \Sigma^*$ we shall denote it by $\delta(q, s)!$.

We identify the following subsets of Σ and Q:

$$\forall q \in Q;\ \Sigma_q := \{\sigma \in \Sigma\ :\ \delta(q, \sigma)!\}, \tag{6}$$

$$\forall q \in Q;\ \forall X \subseteq Q;\ \Sigma_{q,X} := \{\sigma \in \Sigma\ :\ \delta(q, \sigma) \in X\}, \tag{7}$$

$$\forall X \subseteq Q;\ \tau X := \{q \in Q\ :\ \exists \sigma \in \Sigma\ s.t.\ \delta(q, \sigma) \in X\}. \tag{8}$$

Σ_q defines the *active* event set at q, i.e., set of events executable at q, τX is the *source* set of $X \subseteq Q$, that is the set of states which enter X in a single step and $\Sigma_{q,X} \subseteq \Sigma_q$ is the active event subset at q which leads into X. For all state subsets $X \subseteq Q$, we shall denote by $\alpha(X)$ the cardinality of X and by X^c, the complement of X. In particular, $\alpha(Q) < \infty$ and $Q^c = \Phi$.

For the DES G the finite (transient) behaviour will be given by the prefix closed language $L(G) = \{w | w \in \Sigma^*$ and $\delta(q_0, w)!\}$. Sometimes, in order to emphasise that this constitutes the event labels of the paths starting at q_0 of the underlying automaton G, we may, instead, call this behaviour $L(G)_{q_0}$. We shall also mean by the natural sequential (or, non-terminating) behaviour (NTB) of G the limit $[L(G)]^\infty$ which will be assumed to be *non-empty* throughout. In the simplest form the Büchii extension of G can be defined as follows (Eilenberg 1974; Ramadge 1989). Let $Q_m \subseteq Q$ be a given state subset. For each $e \in [L(G)]^\infty$, $\exists$ a unique $S_e: N \longrightarrow Q$, such that $S_e(j) = \delta(q_0, e^j)$, $j \in N$. The state sequence S_e is called the *state trajectory* for the behaviour sequence e. The sequence e and the state trajectory S_e are said to be *admissible*, relative to Q_m (or, simply admissible) if S_e visits the state subset Q_m i.o., i.e., if $S_e(j) \in Q_m$ for infinitely many $j \in N$. The Büchii set Q_m then defines the *acceptable* sequential behaviour of the DES G as follows:

$$[L(G)]^\infty{}_{Q_m} = \{e \mid e\ (\in [L(G)]^\infty)\ \text{and}\ S_e\ \text{are admissible}\}. \tag{9}$$

In principle, one can use multiple Büchii (Eilenberg 1974) sets to generate very complex behaviour. In the sequel, however, only one set will be used. Furthermore, it is obvious that $[L(G)]^\infty$ itself can be taken as the Büchii behaviour of G for $Q_m = Q$.

The control mechanism on DES $G = (Q, \Sigma, \delta, q_0)$ is defined as follows (Ramadge & Wonham 1989): on Σ a partition $\Sigma = \Sigma_c \cup \Sigma_u$ is defined, where the events in Σ_c are called the *controllable* events and those in Σ_u the *uncontrollable* events. Let $\Gamma := \{0, 1\}^{\Sigma_c}$ be the set of binary assignments on Σ_c, i.e., the discrete topology on Σ_c. Then each $\gamma \in \Gamma$, $(\gamma : \Sigma_c \longrightarrow \{0, 1\})$ is a control pattern, in the sense that we say $\sigma \in \Sigma_c$ is *enabled* if $\gamma(\sigma) = 1$ and is *disabled* otherwise. Clearly $\gamma \in \Gamma$ can be extended to whole of Σ by defining $\gamma(\sigma) = 1$, $\forall \sigma \in \Sigma_u$. Given a DES $G = (Q, \Sigma, \delta, q_0)$, the *controlled* DES $G_c := (Q, \Sigma, \delta_c, q_0)$ is defined as follows:

$$\delta_c : \Gamma \times Q \times \Sigma^* \longrightarrow Q\ \text{s.t.},$$
$$(\gamma, q, \sigma) \mapsto \begin{cases} \delta(q, \sigma), & \text{if } (q, \sigma)! \text{ and } \gamma(\sigma) = 1, \\ \text{undefined}, & \text{otherwise} \end{cases} \tag{10}$$

Throughout this paper we shall assume that the Σ partition is *given* and the DES admits the above mentioned control mechanism. We shall, henceforth, drop the subscript c and

assume $G = (Q, \Sigma, \delta, q_0)$ to be a controlled DES, or a controlled discrete-event process (CDEP).

2.3 *Supervisors for CDEPs*

There are two main supervisory control schemes for CDEPs: (a) dynamic (output) feedback, and (b) state-feedback supervisors.

2.3a *Dynamic supervisors:* A dynamic supervisor (Ramadge & Wonham 1987) for a CDEP G is a function $M : L(G) \longrightarrow \Gamma$, which specifies for each string $s \in L(G)$ the subset $\Sigma_s \subseteq \Sigma$ of enabled events ($\Sigma_s := \{\sigma : M(s)\sigma = 1\}$). The closed-loop behaviour of G supervised by M, denoted by $L(G/M)$, is given by:

$$\begin{aligned} &\epsilon \in L(G/M), \text{ and} \\ &\omega\sigma \in L(G/M) \Leftrightarrow \begin{cases} (i) & \omega \in L(G/M), \\ (ii) & M(\omega)\sigma = 1, \textit{ and} \\ (iii) & \omega\sigma \in L(G). \end{cases} \end{aligned} \tag{11}$$

The non-terminating (or, steady-state) behaviour of G supervised by M is, obviously, then given by $[L(G/M)]^{\infty}$.

2.3b *State-feedback supervisors*: More important, from the point of view of applicability and design simplicity, is the state-feedback supervisor (Kumar *et al* 1993a) which, for a given CDEP G is defined as a function $R : Q \longrightarrow \Gamma$. Such a supervisor specifies, for each state of CDEP, the set of enabled events. The closed-loop system, where G is supervised by R, can be viewed as another state machine $(G/R) := (Q, \Sigma, \delta_R, q_0)$, where $\delta_R(q, \sigma) = \delta(q, \sigma)$ if $R(q)\sigma = 1$, and, $\delta_R(q, \sigma)$ is undefined if $R(q)\sigma = 0$. The non-terminating behaviour of the closed-loop is, once again, given by $[L(G/R)]^{\infty}$.

An important property that any supervisor should satisfy is that every response sequence begun (in the closed-loop situation) must be allowed to reach a steady state; in other words, every finite string generated by the closed loop must lead to a non-terminating behaviour sequence. This is called the property of *non-blocking* or the *liveness* property of the system which implies that the system does not stall along a response trajectory. Formally, a supervisor M (dynamic or static) for a CDEP G is said to be *non-blocking* if:

$$\forall s \in L(G/M);\ \exists e \in [L(G/M)]^{\infty} \text{ s.t. } s = e^j \text{ for some } j \in N. \tag{12}$$

Given a CDEP G, and an $X \subseteq Q$, a supervisor M is said to be *X-invariant* if:

$$\forall s \in L(G/M),\ \delta(q_0, s) \in X. \tag{13}$$

An *X-invariant* (or, equivalently, a *state-invariant*) supervisor will, therefore, restrict the

$$[L(G/M)]^{\infty} \neq \Phi, \text{ and}$$
$$[L(G/M)]^{\infty} \subseteq [L(G)]^{\infty}{}_{Q_m}. \tag{14}$$

The supervisor M for G is such that the NTB is non-empty and the set Q_m is visited i.o. Note that this does not imply that the only states visited i.o. are the Q_m states. Also note that this is identical, in principle, to the stability notion of Özveren & Willsky (1991), as introduced earlier, although the setting here is Büchii in contrast, wherein the sequential behaviour is defined in terms of the *admissibility* conditions as given in (9).

The problem addressed to in the sequel is the following:

Given a CDEP $G = (Q, \Sigma, \delta, q_0)$ and the partition $\Sigma = \Sigma_u \cup \Sigma_c$, and a state subset $Q_m \subseteq Q$;

- Determine if there exists a supervisor M for G that is both *non-blocking* and Q_m-*regenerative*. (Henceforth, we shall simply call these supervisors *regenerative* once the set Q_m is understood as identified).
- If M_α is the non-empty family of non-blocking regenerative supervisors then find $M \in M_\alpha$ such that:

$$[L(G/S)]^{\infty} \subseteq [L(G/M)]^{\infty} \quad \forall S \in M_\alpha, \tag{15}$$

i.e., find the minimally restrictive of all the non-blocking supervisors for G.

We shall call these supervisors simply regenerative whenever the Büchii set Q_m is specified or obvious.

3. Main results

In this section we develop a controller synthesis for solving the above mentioned *stabilization* problem. The synthesis procedure is based on construction of a series of state subsets, which are recursively reachable and within the union of which the state invariance of a supervisor can be easily achieved.

3.1 *Control subsets*

Given $G = (Q, \Sigma, \delta, q_0)$, $\Sigma = \Sigma_u \cup \Sigma_c$, and $Q_m (\neq \Phi) \subseteq Q$, define the state subset sequences Q_i and X_i as follows:

$$Q_0 \doteq X_0 \doteq Q_m, \tag{16}$$
$$X_{i+1} \doteq \{q \in (Q_i{}^c \cap \tau Q_i) : \Sigma_{q, Q_i^c} \subseteq \Sigma_c\}, \tag{17}$$
$$Q_{i+1} \doteq Q_i \cup X_{i+1}. \tag{18}$$

Note that X_1, for example, is that subset of Q_0^c which can be driven to Q_0 in a single transition. This is because there *is* a transition to Q_0, and all transitions to Q_0^c are controllable

(and, therefore, can be disabled). Since $\{Q_i\}$ is an ascending chain we define, First Control Subset (FCS) $\hat{Q}$ as its limit, given by:

$$\hat{Q} \doteq \sup_i Q_i \doteq \bigcup_i Q_i \doteq \bigcup_i X_i \tag{19}$$

The FCS $\hat{Q}$ is, therefore, the largest state subset which *reaches* Q_m, possibly uncontrollably, but whose dynamics outside of itself, being *controllable*, can be disabled. To control escape dynamics from Q_m itself, we define, First Regenerative Set (FRS) $\hat{Q}_m$ as:

$$\hat{Q}_m \doteq \{q \in (Q_m \cap \tau\hat{Q}) : \Sigma_{q,\hat{Q}^c} \subseteq \Sigma_c\}. \tag{20}$$

The state set $\hat{Q}_m$ is clearly the Büchii subset which can be made invariant in the closed loop by disabling the exit transitions.

Some interesting properties of the sets X_i, $\hat{Q}$ and $\hat{Q}_m$, which essentially identify controllable transitions at various levels, are listed below (see Valavi for more properties and proofs):

$$\exists k \leq \alpha(Q_m^c) \text{ s.t. } X_i = \Phi \ \ \forall i > k, \text{ and,}$$
$$\forall i \leq k, \ X_i \neq \Phi, \text{ and}$$
$$Q_k = \hat{Q} = \bigcup_{i=0}^{k} X_i. \tag{21}$$

$$\forall i \geq 1, \ \forall q \in X_i, \ \bigcup_{j \geq i} \Sigma_{q,X_j} \subseteq \Sigma_c. \tag{22}$$

$$\forall q \in \hat{Q} - Q_0, \ \Sigma_{q,\hat{Q}^c} \subseteq \Sigma_c. \tag{23}$$

$$\forall i \geq 1, \forall q \in X_i, \ \bigcup_{j=0}^{i-1} \Sigma_{q,X_j} \neq \Phi. \tag{24}$$

Intuitively, these properties all stem from the structure of the X_i chain, which is such that all dynamics from the lower to higher index can be disabled. Therefore, as the first approximation, the supremum of the chain contains all the states that can be driven to Q_m either uncontrollably, or by enabling all controllable events along the path. Also note that property 21 is valid for a finite Q, whereas others hold independently of its cardinality. We shall, however, consider our CDEPs to be finite state only, as earlier mentioned, for practical reasons, even if some of the results hold in general.

Next, we develop the *regenerative* supervisors for our problem as defined in §2.3. This will be achieved by synthesis of a $\hat{Q}$-invariant supervisor, so that all potential paths that escape Q_m indefinitely are disabled. As will be clear from the sequel, it is always possible to do so under the fairly mild condition that q_0 be in $\hat{Q}$, i.e., in one of the X_i's, which are state sets that can always be driven to Q_m in a finite number of steps. Initially, to to keep the synthesis simple and tractable, we shall assume that the FRS $\hat{Q}_m$ equals Q_m. Later this condition will be relaxed.

Note that $\forall q \in Q_m - \hat{Q}_m$; either $\Sigma_{q,\hat{Q}} = \Phi$ or $\Sigma_{q,\hat{Q}^c}$ is uncontrollable. For regenerative behaviour, it should be possible to disable transitions to $\hat{Q}^c$, and for non-blocking condition

there should be a transition to $\hat{Q}$. Therefore, $Q_m - \hat{Q}_m$ is the set of states that violates one or both of these conditions. After the supervisory syntheses given below this problem will be addressed to, so that the final synthesis is free of this apparently severe constraint that Q_m equal $\hat{Q}_m$.

Theorem 1. [*(The R-supervisor)*] *Given the CDEP* $G = (Q, \Sigma, \delta, q_0)$, *with* $\Sigma = \Sigma_u \cup \Sigma_c$, $Q_m \subseteq Q$, *the FCS* $\hat{Q}$ *and the FRS* $\hat{Q}_m$, *assume that* $\hat{Q}_m = Q_m$. *Then there exists a non-blocking and regenerative supervisor R for G if and only if* $q_0 \in \hat{Q}$.

Proof. Sufficiency: Let $q_0 \in \hat{Q}$. Then from the above properties we know that $\hat{Q} = \bigcup_{i=0}^{k} X_i$ for some $k \leq \alpha(Q_m^c) \leq \alpha(Q)$ and $X_i \neq \Phi$, $0 \leq i \leq k$.

Define a state feedback supervisor R as follows:

$$\forall q \in Q_0,\ R(q)\sigma = 0\ \forall \sigma \in \Sigma_{q,\hat{Q}^c} \quad [\subseteq \Sigma_c;\ (20)], \tag{25}$$

$$\forall i \geq 1, \forall q \in X_i,\ R(q)\sigma = 0\ \forall \sigma \in (\textstyle\bigcup_{j \geq i}^{k} \Sigma_{q,X_j}) \cup (\Sigma_{q,\hat{Q}^c}), \quad [\subseteq \Sigma_c;\ \text{properties 21, 23}] \tag{26}$$

$$R(q)\sigma = 1, \text{otherwise.} \tag{27}$$

R, as is obvious, is a supervisor that drives the system fastest to Q_m, the desired state set for regenerative operation, by preventing escape from $\hat{Q}$ as well as any backward movement (towards increasing index on Q_i) within $\hat{Q}$. Therefore, once the system is in $\hat{Q}$, it will move into Q_m in at most k steps.

The closed loop system $G/R \doteq (Q, \Sigma, \delta_R, q_0)$, is then defined by:

$$\delta_R(q,\sigma) = \begin{cases} \delta(q,\sigma), & \text{if } R(q)\sigma = 1, \text{and} \\ \text{undefined}, & \text{if } R(q)\sigma = 0 \end{cases} \tag{28}$$

To show that R is *non-blocking*, note that $\forall q \in \hat{Q}$, $\delta_R(q,\sigma)!$ (is defined) $\Longrightarrow \delta_R(q,\sigma) \in \hat{Q}$, because transitions from $\hat{Q}$ to its complement are disabled by the supervisor R ((25), (26)). Therefore, it follows that:

$$q_0 \in \hat{Q} \Longrightarrow \delta_R(q_0, s) \in \hat{Q},\ \forall s \in L(G/R). \tag{29}$$

In other words, R is a $\hat{Q}$-invariant supervisor. Furthermore, (20), (24), (27) and the fact that $\hat{Q} = Q_0 \cup (\bigcup_{i=1}^{k} X_i)$ together imply that:

$$\forall q \in \hat{Q};\ \{\sigma \in \Sigma_{q,\hat{Q}} : R(q)\sigma = 1\} \neq \Phi. \tag{30}$$

Next, from (29) and (30), it follows that $\forall s \in L(G/R)$, $\exists \sigma \in \Sigma$ such that $s\sigma \in L(G/R)$. Therefore, $\forall s_0 \in L(G/R)$, $\exists$ a sequence $s_0 \prec s_1 \prec s_2 \cdots$ in $L(G/R)$, so that R is a non-blocking supervisor that keeps the closed-loop live.

To show that R is *regenerative*, it suffices to show that a transition to $X_0 (= Q_m)$ is guaranteed in at most k steps. To this end assume that $s, s\omega \in L(G/R)$, with $\delta_R(q_0, s) = q \in X_r$, for some $1 \leq r \leq k$ (this is guaranteed by $q_0 \in \hat{Q}$ and the supervisor design (25)–(27), and $|\omega| \geq k$. Let $\sigma_{k-1}\sigma_{k-2}\cdots\sigma_0$ be the k-long prefix of ω (note the reverse index on ω). Clearly $\bar{q}_j = \delta_R(\bar{q}_{j-1}, \sigma_{k-j}) \in X_{r_j}$, where $\bar{q}_0 = q$ and $r > r_1 > r_2 \cdots$.

Since the index r_i decreases at least as fast as the index on σ, we have $q_i \in X_0$ for some $i \leq k$. Furthermore, if $q \in X_0$, $\delta_R(q, \sigma_{k-1}) \in X_0 \cup X_r$, for some $1 \leq r \leq k$, the same process holds. Thus, in the closed loop a transition to $Q_m (= X_0)$ is always guaranteed (in $\leq k$ steps) and, therefore, the supervisor is regenerative.

Necessity. See appendix A. □

Remark 1. It is obvious from appendix A that *regenerative* and *non-blocking* supervisors are $\hat{Q}$-invariant, since for every *sub*-CDEP that evolves from a reachable state $q \in Q$, we have that $q \in Q_i$ (for some i) $\subseteq \hat{Q}$. This implies that only the states from $\hat{Q}$ can be regeneratively driven to Q_m along non-blocking trajectories.

3.1a *Minimally restrictive supervisors:* The R-supervisor, though *non-blocking* and *regenerative* is, very restrictive in terms of the richness of the closed-loop behaviour. In fact, as is obvious, the R-action is to select control patterns to drive the closed-loop to Q_m as fast as possible. This constrains the behaviour profile of the system. To overcome this problem we use another supervisor in parallel with the R-supervisor, so that the closed-loop is both rich in behaviour (*minimally restrictive*) as well as regenerative.

Define a state-feedback supervisor $S : Q \longrightarrow \Gamma$ for the CDEP G as:

$$\forall q \in \hat{Q};\ S(q)\sigma = 0\ \ \forall \sigma \in \Sigma_{q,\hat{Q}^c}\ [\subseteq \Sigma_c], \tag{31}$$

$$S(q)\sigma = 1\ \text{ otherwise.} \tag{32}$$

If $q_0 \in \hat{Q}$ the supervisor S is $\hat{Q}$-invariant, since $\forall q \in \hat{Q}$ the transitions that lead to a state in $\hat{Q}^c$ are disabled. Note that the supervisor S, while ensuring that the system remain inside $\hat{Q}$, does not drive it hard to Q_m, the way R does. We now have the following straightforward result.

Theorem 2. *Let G be a CDEP, as in theorem 1, with $\hat{Q}_m = Q_m$, and $q_0 \in \hat{Q}$. Let S be the state feedback supervisor for G as defined above. Then, if M is any non-blocking and regenerative supervisor for G, we have that:*

$$L(G/M) \subseteq L(G/S). \tag{33}$$

In other words, S is the minimally restrictive $\hat{Q}$-invariant supervisor, richer in behaviour than the non-blocking and regenerative supervisor subclass.

Proof. (Straightforward) We know that every *non-blocking* and *regenerative* supervisor is $\hat{Q}$-invariant under the given conditions (see remark 1). Therefore, since $q_0 \in \hat{Q}$ and the states in $\hat{Q}^c$ cannot be driven to Q_m in the desired manner, every regenerative and non-blocking supervisor must disable transitions to $\hat{Q}^c$.

Therefore, $s \in L(G/M) \Longrightarrow \delta(q_0, \omega) \in \hat{Q},\ \forall \omega \in pr(s)$. Also since the only transitions that S disables are those to $\hat{Q}^c$, we have that $\delta(q_0, \omega) \in \hat{Q} \Longrightarrow \omega \in L(G/S),\ \forall \omega \in pr(s) \Longrightarrow s \in L(G/S)$. The empty string ϵ in $pr(s)$ does not affect the argument since $q_0 \in \hat{Q}$. This completes the proof. □

From the preceeding discussion it is obvious that the R-supervisor capability of fastest motion to the desired state set constrains its behaviour, whereas, in contradistinction, the S-supervisor's minimally restrictive behaviour limits its frquency of arrival into the desired state set. In practice, a compromise is desirable. In the rest of this section we shall address this problem and provide a solution based on continual switching between R and S supervisors that, dynamically, takes into account the last arrival into Q_m.

First we define a *delay* function $g : L(G) \longrightarrow N \cup \{0\}$ (N is the set of natural numbers) as:

$$g(\epsilon) = 0,$$
$$g(s\sigma) = \begin{cases} g(s) + 1 & \text{if } \delta(q_0, s\sigma) \in Q_m^c, \text{ and} \\ 0 & \text{if } \delta(q_0, s\sigma) \in Q_m. \end{cases} \tag{34}$$

The *delay* function is aperiodically cyclic and in every cycle counts the number of events executed outside of Q_m, so that a large value in any cycle reflects large behaviour executed without returning to Q_m. Next, given a $k \in N$, we define a **k-jump** function $f_{\mathbf{k}} : N \cup \{0\} \longrightarrow \{0, 1\}$ as:

$$f_{\mathbf{k}}(i) = \begin{cases} 0, & \text{if } i < k, \text{ and} \\ 1, & \text{if } i \geq k. \end{cases} \tag{35}$$

The composite sequence $L(G) \xrightarrow{g} N \cup \{0\} \xrightarrow{f_k} \{0, 1\}$, represented by the function composition $f_{\mathbf{k}} \circ g$, therefore, becomes *high* every time the CDEP executes more than **k** events in a row outside Q_m. The composite function $f_{\mathbf{k}} \circ g$ will henceforth be called the system **switching function** of order k.

The *switching function* is used for control strategy as follows: Run the CDEP G under the supervisor S as long as the switching function is *low* (= 0) and once the switching function becomes *high* (= 1) the supervisor R takes over. The parameter k can be dynamically varied to satisfy the requirement on the rate of visitations to Q_m. Note that $k = 0$ implies running the system only under R, since the fastest convergence to Q_m is desired and as $k \to \infty$ only S will be used since the constraint now is the richness of behaviour rather than fast settling to Q_m. The composite overall dynamic supervisor (see 2.3a) T modelling the abovementioned control strategy can be defined as under.

DEFINITION. Given a CDEP $G = (Q, \Sigma, \delta, q_0)$, $\Sigma = \Sigma_u \cup \Sigma_c$ and the associated state sets Q_m, $\hat{Q}$, and $\hat{Q}_m$ such that $\hat{Q}_m = Q_m$ and $q_0 \in \hat{Q}$. Let $k \in N \cup \{0\}$ be given and the supervisors R and S be as defined earlier. Further let the switching function $f_{\mathbf{k}} \circ g$ be as defined above.

Define the dynamic supervisor $T : L(G) \longrightarrow \Gamma$ as:

$$T(s)\sigma = \{[f_{\mathbf{k}} \circ g(s)] \wedge [R(\delta(q_0, s))\sigma]\} \vee \{\neg[f_{\mathbf{k}} \circ g(s)] \wedge [S(\delta(q_0, s))\sigma]\}. \tag{36}$$

From the preceeding discussion it follows that for finite k the supervisor T is both *non-blocking* and *regenerative* with the richness of behaviour increasing as k becomes

ny specification changes or needs of the environment.

The computational effort involved in T synthesis is mainly in the recursive computation f the sets X_i , so that a loose (conservative) upper bound on the complexity is $\bigcirc(n^2 p)$, here $n = \alpha(Q)$ and p is the number of edges of G.

.3 *Weak supervisory conditions*

As explained in § 3.1, for regenerative behaviour it should be possible to disable transitions o $\hat{Q}^c$ and for non-blocking behaviour transitions to $\hat{Q}$ must be ensured. The states in $Q_m - \hat{Q}_m$ are precisely those Büchii states that violate one or both of these conditions. Therefore, for desired behaviour it must be ensured that these states are kept out of the ossible closed-loop paths with minimal reduction in the desired behaviour set.

This can be done as follows; Let G be given such that $\hat{Q}_m \neq Q_m$. We proceed to build family of control subsets as follows.

et $Q_{m_0} \doteq Q_m$; $\hat{Q}_0 \doteq \hat{Q}$; $\hat{Q}_{m_0} \doteq \hat{Q}_m$. Define:

$$Q_{m_i} = \hat{Q}_{m_{i-1}}, \tag{37}$$

$$\hat{Q}_i = \text{set obtained from FCS construction with } Q_{m_i} \text{ as } Q_m, \tag{38}$$

$$\hat{Q}_{m_i} = \text{set obtained from eqn. 20 with } Q_{m_i} \textit{ as } Q_m \text{ and } \hat{Q}_i \text{ as } \hat{Q} \tag{39}$$

Note that the chains Q_{m_i}, $\hat{Q}_i$ and $\hat{Q}_{m_i}$ are all *descending*. Next define:

$$\hat{Q}_s \doteq \inf_i \hat{Q}_i \doteq \bigcap_i \hat{Q}_i, \tag{40}$$

$$Q_{m_s} \doteq \inf_i Q_{m_i} \doteq \bigcap_i Q_{m_i}, \tag{41}$$

$$\hat{Q}_{m_s} \doteq \inf_i Q_{m_i} \doteq \bigcap_i \hat{Q}_{m_i}. \tag{42}$$

The role played by $\hat{Q}_s$ and $\hat{Q}_{m_s}$ is the same as that of $\hat{Q}$ and $\hat{Q}_m$, respectively, in the earier case when $\hat{Q}$ was assumed equal to $\hat{Q}_m$. In other words, transitions to $\hat{Q}_{m_s}$ can always e guaranteed from $\hat{Q}_s$, whereas from outside of $\hat{Q}_s$ this cannot be achieved. Therefore, ve shall call $\hat{Q}_s$ the Infimal control subset (ICS) and $\hat{Q}_{m_s}$ the Infimal regenerative subset IRS). It is easily verified that $\hat{Q}_{m_s} = Q_{m_s}$. We call the set Q_{m_s} the Infimal marked subset IMS). It is obvious that if the algorithm of theorem 1 is run using $\hat{Q}_m$ as Q_m, $\hat{Q}_s$ *as* $\hat{Q}$ nd $\hat{Q}_{m_s}$ *as* $\hat{Q}_m$, then the following generalization of that theorem holds.

Theorem 3. (Generalization) *Given a CDEP* $G = (Q, \Sigma, \delta, q_0)$, *with* $\Sigma = \Sigma_u \cup \Sigma_c$, *nd the Büchii state subset* $Q_m \subseteq Q$, *let* $\hat{Q}_s$ *and* $\hat{Q}_{m_s}$ *be the associated ICS and IRS, espectively. Then G is a regenerative CDEP (i.e., there exists a regenerative supervisor or G) if and only if* $q_0 \in \hat{Q}_s$.

The example given in appendix B highlights the advantages of the synthesis given here relative to the well-known supervisory synthesis of Ramadge (1989). The principal feature in favour of synthesis in by Ramadge (1989) is that the NTB specifications are given in terms of subsets of $[L(G)]^{\infty}$, which is more general than specifying i.o. visitations to a given Büchii state subset. However, the approach requires conversion of the NTB to finite behaviour using the prefix operator $pr(.)$ and then synthesising the supervisor that meets the prefix specifications. The limitations of this method are as below.

- A supervisor having a finite state structre is possible only if $pr(B)$ is regular (Ushio 1993).
- The synthesis requires that B be closed relative to $[L(G)]^{\infty}$, which may be hard to obtain in practice (see appendix B). In case B does not satisfy this constraint, the closed-loop can only be guaranteed to stay within $\bar{B}$, the topological closure of B relative to $[L(G)]^{\infty}$, and not in B which is the NTB specification (see appendix B).

The acceptable specifications used here, though relatively less general, are meaningful and adequate for a large CDEP class. Main advantages of our method, for example, are the following.

- The Büchii set specification is very meaningful in terms of closed -loop stability analysis (Özveren & Willsky 1991). In fact, stability specifications in terms of Büchii sets is a very natural and logical extension of CDEPs based on finite state machine models (Eilenberg 1974). The synthesis then provides *flexible stabilization* of CDEPs.
- The method does not require construction of a grammar for acceptable NTB, thus reducing the computational complexity.
- It is easy to deal with, possibly conflicting, finite and non-terminating behaviour. This is due to the supervisor structure obtained and the operational flexibility achieved due to our *switching function.*

The supervisory synthesis approach as outlined above, therefore, presents a very efficient mechanism of ensuring flexible stabilization and satisfaction of other closed-loop specifications of CDEPs.

The authors are deeply grateful to one of the anonymous reviewers for pointing out a serious error in the proof of theorem 1 as well as other minor errors. They are also thankful to other reviewers for many helpful suggestions.

Appendix A

Completion of the proof of theorem 1 requires some auxillary definitions and lemmas. We begin by introducing the notion of a *restricted* supervisor. Let $G = (Q, \Sigma, \delta, q_0)$ be a CDEP. Suppose $\delta(q_0, s) = \acute{q} \in Q$ for some $s \in \Sigma^*$. Then $\acute{G} = (\acute{Q}, \Sigma, \acute{q}, \acute{\delta})$, where

Büchii set for the *sub*-CDEP will be given by $\acute{Q}_m = \acute{Q} \cap Q_m$.

DEFINITION. Let G be a CDEP and $\acute{G}$ a *sub*-CDEP of G. Suppose $T : Q \longrightarrow \Gamma$ is a supervisor for G. Then $\acute{T} : \acute{Q} \longrightarrow \Gamma$ is said to be the *restriction* of T to $\acute{Q}$ if:

$$\acute{T}(q)\sigma = T(q)\sigma \quad \forall q \in \acute{Q},\ \forall \sigma \in \Sigma. \tag{A1}$$

It follows directly from the above definition and the notions of *non-blocking* and *regenerative* supervision introduced earlier that if T is a *non-blocking* and *regenerative* supervisor for G, so is $\acute{T}$ for $\acute{G}$. Next we explore the structure of X_i , Q_i and $\hat{Q}$ ((16)–(18)) sets of G when the Büchii set is restrictively altered.

Lemma A1. Let G be a CDEP with Q_m its associated Büchii set. Let $\bar{Q}_m \subseteq Q_m$, and let $\bar{Q}_i$, $\bar{X}_i$ and $\hat{\bar{Q}}$ be the control subsets of G relative to Büchii set $\bar{Q}_m$. Then $\bar{Q}_i \subseteq Q_i$, $\forall i$.

Proof (By induction on i). At $i = 0$, we have that $\bar{X}_0 = \bar{Q}_0 = \bar{Q}_m$. Now $Q_0 = Q_m \Longrightarrow \bar{Q}_0 \subseteq Q_0$ since $\bar{Q}_m \subseteq Q_m$.

Assume that the hypothesis holds for $i = j$.

Now by definition (see (17)), we have that $q \in \bar{X}_{j+1} \Longrightarrow q \in (\tau \bar{Q}_j \cap \bar{Q}_j^c)$ and $\Sigma_{q,\bar{Q}_j^c} \subseteq \Sigma_c$. Therefore $q \in \tau Q_j \cap (Q_j^c \cup (Q_j - \bar{Q}_j)) = (\tau Q_j \cap Q_j^c) \cup (\tau Q_j \cap (Q_j - \bar{Q}_j))$ because $\bar{Q}_j \subseteq Q_j$ and $\bar{Q}_j^c = Q_j^c \cup (Q_j - \bar{Q}_j)$. Also note that $\Sigma_{q,\bar{Q}_j^c} \subseteq \Sigma_c \Longrightarrow \Sigma_{q,Q_j^c} \subseteq \Sigma_c$ since $Q_j^c \subseteq \bar{Q}_j^c$.

We, therefore, have that $\bar{X}_{j+1} = X_{j+1} \cup \{q \in \tau Q_j \cap (Q_j - \bar{Q}_j) : \Sigma_{q,Q_j^c} \subseteq \Sigma_c\}$, where the second set is obviously in Q_j. Therefore, $\bar{X}_{j+1} \subseteq X_{j+1} \cup Q_j = Q_{j+1}$, and, consequently, $\bar{Q}_{j+1} = \bar{Q}_j \cup \bar{X}_{j+1} \subseteq Q_j \cup X_{j+1} = Q_{j+1}$. Hence $\bar{Q}_j \subseteq Q_j, \forall j \in N$. □

Remark A1. It is obvious from the above proof that lemma A1 holds if G is replaced by a *sub*-CDEP and the Büchii subset is compatibly chosen, as discussed earlier.

Next the proof of the necessity of theorem 1 follows:

Proof (Necessity of theorem 1). Let $G = (Q, \Sigma, \delta, q_0)$ be the given CDEP with $Q_m (\neq \Phi)$ as its associated Büchii state subset. Let T be a *non*-blocking and *regenerative* supervisor for G.

We shall prove the necessity of the theorem 1 by induction on $\alpha(Q)$.

For $\alpha(Q) = 1$, since $q_0 = Q_m$, the result is obvious. Assume that the result holds for $\alpha(Q) = n$. Next, let $\alpha(Q) = n + 1$. Suppose $\exists \sigma \in \Sigma$ such that $\delta(q_0, \sigma) = q_0$. Then if $\sigma \in \Sigma_u$, we shall have $q_0 \in Q_m$, otherwise no *regenerative* supervisor can exist as the CDEP could indefinitely stay in q_0. If this were indeed the case then $q_0 \in \hat{Q}$ and there is nothing further to prove. Therefore, without loss of any generality, we shall assume that in such a case $\sigma \in \Sigma_c$.

Next let $q_1, q_2, \cdots, q_r; r \leq n$ be the states, other than q_0, which are 1-step accessible from q_0, i.e., $\forall i \leq r$, $q_i = \delta(q_0, \sigma_i)$ for some $\sigma_i \in \Sigma$. Consider the *sub*-CDEPs $\acute{G}_i$, $1 \leq$

$i \leq r$, generated from the states q_i, i.e., $\acute{G}_i = (\acute{Q}_i, \Sigma, q_i, \acute{\delta})$, where $\acute{Q}_i \subseteq Q$ is the state set accessible from q_i, and $\acute{Q}_{im} = Q_m \cap \acute{Q}_i$ is the associated Büchii set. Since $\alpha(\acute{Q}_i) \leq n$, we have, from lemma A1 and remark A1, that, $\forall i, j; i \leq r$ and $j \leq r_i$, $\acute{Q}_{ij} \subseteq Q_j$, i.e., the jth control subset corresponding to the ith *sub*-CDEP is contained in the jth control subset of G. Here r_i denotes the length of the control subset chain upto its saturation (see (21)).

Furthermore, since the restricted supervisors $\acute{T}_i$ acting on $\acute{G}_i$ are *non-blocking* and *regenerative*, and since $\alpha(\acute{Q}_i) \leq n$, it follows by induction hypothesis that $q_i \in \hat{\acute{Q}}_i (= \acute{Q}_{ir_i}$, by (19)), $\forall i$. Next let $\hat{r} = \max_i r_i$. Then it follows that $\{q_1, q_2, \cdots, q_r\} \subseteq Q_{\hat{r}}$. Clearly then, $q_0 \in \tau Q_{\hat{r}} \cap Q_{\hat{r}}^c$ and $\Sigma_{q_0, Q_{\hat{r}}^c} \subseteq \Sigma_c$ (since the only such transition can be to q_0 itself and that is controllable, by assumption); so that $q_0 \in Q_{\hat{r}+1} \subseteq \hat{Q}$. This completes the proof. □

5. A simple synthesis example

Consider the CDEP given below in the figure. Here $Q = \{q_0, q_1, q_m\}$, $\Sigma = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ with $\Sigma_c = \{\sigma_2\}$ (say). Clearly $L(G) = \sigma_1\sigma_2^*(1 + \sigma_3\sigma_4^*)$ and $[L(G)]^\infty = \{\sigma_1\sigma_2^\omega, \sigma_1\sigma_2^*\sigma_3\sigma_4^\omega\}$. Let the acceptable (regenerative) non-terminating behaviour (NTB) be infinitely often i.o. visitations to the set $Q_m \doteq \{q_m\}$. The corresponding largest desired behaviour then is $B = \{\sigma_1\sigma_2^*\sigma_3\sigma_4^\omega\} \subseteq [L(G)]^\infty$. With this as the desired NTB, the synthesis procedure of [*R*] gives a supervisor M, whose finite and NTB behaviour equals $L(G/M) = \sigma_1\sigma_2^*(1 + \sigma_3\sigma_4^*)$ and $[L(G/M)]^\infty = \{\sigma_1\sigma_2^\omega, \sigma_1\sigma_2^*\sigma_3\sigma_4^\omega\}$. The closed-loop NTB obtained is actually the (topological) closure (see §2.1) of B relative to $[L(G)]^\infty$ and violates the specification of the desired NTB. This limitation of the procedure occurs because B is not closed relative to $[L(G)]^\infty$.

If however we choose $B = \{\sigma_1\sigma_3\sigma_4^\omega\}$, which is closed relative to $[L(G)]^\infty$, as the desired NTB, then the algorithm of by Ramadge (1989) results in a supervisor M such that $L(G/M) = \{\sigma_1\sigma_3\sigma_4^*\}$ and $[L(G/M)]^\infty = \{\sigma_1\sigma_3\sigma 4^\omega\}$ which satisfies the NTB specification. Note, however, that the resulting supervisor turns out to be very restrictive.

To obtain a less restrictive supervisor, we could choose $B_k = \{\sigma_1\sigma_2^k\sigma_3\sigma_4^\omega\}$, $k < \infty$, which is closed w.r.t. $[L(G)]^\infty$. With this B the synthesis by Ramadge (1989) results in a supervisor less restrictive than the earlier one (larger k improves the performance), but the procedure requires construction of a grammar for $pr(B_k)$ for every desired k.

Using the algorithm of §3, we obtain the following contol subsets as follows.

$X_0 = Q_m = \{q_m\}$; $X_1 = \{q_1\}$; $X_2 = \{q_0\}$. $\hat{Q}_s = Q$; $\hat{Q}_{m_s} = Q_m$.

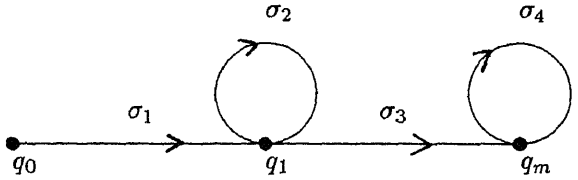


Figure B1. A CDEP example.

Observe that $q_0 \in \hat{Q}_s$. We, therefore, get:
The supervisor $R \doteq R(q)\sigma = 0$ if $q = q_1$ and $\sigma = \sigma_2$; and $R(q)\sigma = 1$ otherwise.
The supervisor $S \doteq S(q)\sigma = 1\ \forall q \in Q$ and $\sigma \in \Sigma$.
The NTB of the closed-loop under the composite supervisor T is then equal to $B_k = \{\sigma_1\sigma_2^k\sigma_3\sigma_4^\omega\}$ which is the desired NTB. The synthesis, however, avoids grammar construction as well as specification of the language that guarantees the desired regenerative behaviour, e.g. $pr(B)$, which is necessary for other approaches. The parameter k here, in fact, parametrises the family of non-blocking regenerative supervisors for the system in the example given.

References

Brave Y, Heyman M 1989 On stabilization of discrete event processes. In *IEEE Proceedings of 28th Conference on Decision and Control*, Tampa, FL, pp 2737–2742

Eilenberg S 1974 *Automata, languages and machines* (New York: Academic Press) vol. A

Hopcroft J E, Ullman J D 1979 *Introduction to automata theory, languages and computation* (Reading, MA: Addison-Wesley)

Özveren C M, Willsky A S 1991 Stability and stabilizability of discrete event dynamic systems. *J. ACM* 38: 730–752

Ramadge P J 1989 Some tractable supervisory control problems for discrete event systems described by Büchii automata. *IEEE Trans. Autom. Control* 34: 10–19

Kumar R, Garg V, Marcus S I 1992 On supervisory control of sequential behaviours. *IEEE Trans. Autom. Control* 37: 1978–1984

Kumar R, Garg V, Marcus S I 1993 Predicates and predicate transformers for supervisory control of discrete event dynamic systems. *IEEE Trans. Autom. Control* 38: 232–247

Kumar R, Garg V, Marcus S I 1993 Language stability and stabilizability of discrete event dynamical systems. *SIAM J. Control Optimization* 31: 1294–1320

Ramadge P J, Wonham W M 1987 Supervisory control of a class of discrete event processes. *SIAM J. Control Optimization* 25: 206–230

Ushio T 1993 A necessary and sufficient condition for the existence of finite state supervisors for discrete event systems. *IEEE Trans. Autom. Control* 38: 135–137

Valavi A G 1994 *A theory for regenerative supervision of discrete event processes.* M Tech thesis, Indian Institute of Technology, Kanpur

Wonham W M, Ramadge P J 1987 On the supremal controllable sublanguage of a given language. *SIAM J. Control Optimization* 25: 637–659

On the behaviour of organised disturbances in a turbulent boundary-layer

P K SEN and SRINIVAS V VEERAVALLI

Department of Applied Mechanics, Indian Institute of Technology,
New Delhi 110 016, India

MS received 29 March 1997; revised 17 July 1997

Abstract. This paper is in continuation of our earlier work on the role of hydrodynamic stability theory in understanding wall-bounded turbulent flows. Work in this area was pioneered by Malkus, followed by Reynolds, Tiederman and Hussain. Numerical results showed that the linear instability modes are damped, a result also confirmed by our earlier work for the boundary layer flow case. This led to waning of interest in this approach.

In the present work the problem is reformulated using an improved non-isotropic model for the stress tensor based on the model of Pope. This improved model does yield unstable wall modes. A wide range of unstable wavenumbers is observed and these unstable modes mimic some of the key features of wall-bounded turbulent flows. Comparisons with experimental data are also presented. The present work keeps alive the question of relevance of stability theory to wall-bounded turbulent flows.

Keywords. Hydrodynamic stability; wall-bounded turbulent flow; boundary layer flow; non-isotropic model.

1. Introduction

The motivation for the present paper is primarily to re-examine a question that has been asked before, but has got mixed answers so far. Whether or not hydrodynamic stability theory has any bearing on the problem of wall-bounded turbulent shear flows, especially in view of the fact that it does have a significant role in free turbulent shear flows. See, for example, the discussion by Gaster *et al* (1985) on the turbulent mixing layer, the comprehensive review on stability and free turbulent shear flows by Liu (1988), and a more recent overview by Roshko (1992).

The question of a connection between stability theory and turbulent shear flow was first raised by Landau (1944) based on a nonlinear stability model. While his work did not prove to be a suitable model for turbulence, his equation for nonlinear growth found

many applications in the field of instability and transition. The next work of great conceptual importance is that of Malkus (1956), wherein a theory of turbulence was developed based on the introduced concept of marginal stability. According to this concept, it was proposed that if the mean velocity profile typical of wall-bounded turbulent flows is used in the solution of the classical Orr–Sommerfeld equation, then the profile would prove to be marginally or neutrally stable at the existing flow Reynolds number. It is significant also that Malkus suggested that the *molecular viscosity* and not the eddy viscosity be used in the solution of the Orr–Sommerfeld equation. Malkus's theory was rigorously put to test in an important work by Reynolds & Tiederman (1967). This work also gives a lucid review of Malkus's theory. Reynolds & Tiederman (1967) investigated the stability of fully developed turbulent flow between parallel plates, on the lines of Malkus's proposed theory. They used the turbulent mean velocity profile for channel flow in the solution of the classical Orr–Sommerfeld equation (using the molecular viscosity in the Orr–Sommerfeld equation). The results obtained showed without doubt that Malkus's theory as proposed was not valid, and there was a huge discrepancy between the flow Reynolds number and the Reynolds number corresponding to neutral or marginal stability, the latter being an order of magnitude higher than the former. Prabhu (1968, private communication) also independently obtained the same result. More recently, Sen *et al* (1993) obtained the same result for the turbulent boundary layer flow problem.

Although Malkus's theory was not proven, it did not immediately detract attention from the general question of connection between instability and turbulence, and, as mentioned before, considerable success was obtained later on in linking the two for the cases of free turbulent shear flows. The results for free turbulent shear flows were very decisive mainly because of the inflectional mean-velocity profile present in such flows. Thus the dominant instability is the *inviscid* instability, which continues to exist even in the presence of the existing turbulent flow field. Also, the dominant eddy structure of the turbulent flow field is found to correlate well with the vorticular structure corresponding to the inviscid instability (see Liu 1988 and Roshko 1992 for comprehensive reviews).

The general question of connection between instability and turbulence, with reference to wall-bounded turbulent shear flow, went through another round of serious examination by Reynolds & Hussain (1972). This time, abandoning Malkusian precepts, first the basic equations were derived for a superposed organised disturbance in turbulent flow. Thereafter an equation for the stability of this organised disturbance was derived, which was like an augmented form of the classical Orr–Sommerfeld equation, but containing extension terms dependent on the *eddy viscosity*. Underlying the model was a closure problem for the Reynolds stress tensor, which was resolved based on what the authors called the 'Newtonian eddy' model. Results of solution of their extended Orr–Sommerfeld equation, for channel flow, again yielded damped modes. Nevertheless, the results were closer to establishing a connection between instability and turbulent wall flows than was obtained in Reynolds & Tiederman (1967). Experiments performed by Hussain & Reynolds (1970, 1972) showed some agreement with the theory of Reynolds & Hussain (1972). Subsequent to the works of Reynolds & Hussain (1972), some nonlinear and three-dimensional theories were also developed (see for example the work of Jang *et al* 1986), but these are outside the scope of the present discussion.

The present work seeks to re-examine the question of connection between instability nd wall turbulence. The organised disturbances are linear and two-dimensional, and the articular flow chosen is the turbulent boundary layer flow over a flat plate. The chief nprovement over the work of Reynolds & Hussain (1972) is that a more realistic and nproved model is chosen for the turbulent stress tensor, based on the anisotropic model f Pope (1975). This model gives further extensions of the Orr–Sommerfeld equation ver what Reynolds & Hussain (1972) obtained. The results show that an unstable wall-node exists over a wide range of the spatial wavenumber α. The instability characteristics cale very well with the inner variables of turbulent flow. Also the organised disturbances imic some of the key features of wall-turbulence. Experiments performed for the one imensional energy-spectrum for the turbulent longitudinal velocity, indicate that the range f unstable wavelengths is well contained within the energy-containing part of the energy pectrum. Thus, the present work keeps alive the question of relevance of stability theory n understanding turbulence in wall-bounded turbulent flows.

Preliminary versions of the present work appeared in Sen & Veeravalli (1994, 1996).

. Theory

he instantaneous velocity vector u_i obeys the Navier-Stokes and continuity equations:

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j}, \tag{1a}$$

$$\frac{\partial u_i}{\partial x_i} = 0. \tag{1b}$$

he velocity and pressure fields are decomposed in turbulent flows by the well known eynolds decomposition; typically,

$$u_i = \overline{u}_i + u_i'; \quad p = \overline{p} + p'. \tag{2}$$

Here $\overline{u}_i$, $\overline{p}$ are respectively the mean velocity and pressure, and u_i', p' are the (random) luctuations. If we now superpose an organised (solenoidal) disturbance $\tilde{u}_i$, $\tilde{p}$ (with zero nean), the instantaneous velocity and pressure are respectively given as follows:

$$u_i = \overline{u}_i + \tilde{u}_i + u_i'; \quad p = \overline{p} + \tilde{p} + p'. \tag{3a}$$

he time averages of u_i, p are still respectively $\overline{u}_i$, $\overline{p}$, but the ensemble (phase-locked) verages are different. The ensemble averages are given respectively as

$$\langle u_i \rangle = \overline{u}_i + \tilde{u}_i; \quad \langle p \rangle = \overline{p} + \tilde{p}. \tag{3b}$$

Moreover, the organised disturbance is assumed to be small, or linear, and in addition it obeys the following:

$$|\langle \tilde{u}_i \tilde{u}_j \rangle| \ll |\langle u'_i u'_j \rangle|. \tag{4}$$

The above assumption (4) restricts the organised disturbances to being weaker than was considered by Reynolds & Hussain (1972).

For future clarity some definitions and notations are introduced: (i) an overbar (–) over any quantity implies time average; (ii) the symbols $\langle\rangle$ enclosing a quantity denote ensemble average. Further, we note that: (i) when the flow field is not perturbed by an organised disturbance, the time average and the ensemble average are the same; (ii) when the flow field contains an organised disturbance, the time average and the ensemble average are different; (iii) it will be seen later that if the organised disturbance in the flow field is small (in the sense defined in (4) above), then the *time average* is the same as the time average of the corresponding flow field that is *not perturbed* by the organised disturbance. Hereafter, the terms 'perturbed flow' and 'unperturbed flow' will mean respectively the flow that is perturbed and that is not perturbed, by the organised disturbance.

One needs now to look into the forms that (1a) takes upon different ways of averaging. First the typical time average, with the disturbance decomposition as in (2), which yields the well-known Reynolds averaged equation:

$$\frac{\partial \overline{u}_i}{\partial t} + \overline{u}_j \frac{\partial \overline{u}_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \overline{p}}{\partial x_i} + \nu \frac{\partial^2 \overline{u}_i}{\partial x_j \partial x_j} - \frac{\partial (\overline{u_i' u_j'})}{\partial x_j}. \tag{5}$$

One can also obtain the ensemble average of (1a), to obtain the equation for $\langle u_i \rangle$. We need to remember the following rules for ensemble averaging, given with respect to any two generic quantities a, b.

$$\langle \overline{a}\tilde{b} \rangle = \overline{a}\tilde{b}; \quad \langle a'\tilde{b} \rangle = 0; \tag{6}$$

and, in view of (4), and in the event that $\tilde{a}$, $\tilde{b} \sim O(\tilde{u}_i)$, then the following is also valid:

$$\langle \tilde{a}\tilde{b} \rangle \approx 0; \quad \overline{\tilde{a}\tilde{b}} \approx 0. \tag{7}$$

In view of (4) and (7), it may be noted that the time average of (1a), even when disturbance form (3a) is being considered, becomes the same as (5). Particularly, the organised disturbances are assumed small enough not to alter $\overline{u_i' u_j'}$. Further discussion on this point will be taken up later.

Next we look at the form of (1a) after ensemble averaging, with the disturbance decomposition as in (3a). Remembering (6) and (7) one obtains:

$$\frac{\partial \langle u_i \rangle}{\partial t} + \langle u_j \rangle \frac{\partial \langle u_i \rangle}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \langle p \rangle}{\partial x_i} + \nu \frac{\partial^2 \langle u_i \rangle}{\partial x_j \partial x_j} - \frac{\partial (\langle u_i' u_j' \rangle)}{\partial x_j}. \tag{8}$$

One is now in a position to obtain the dynamic equation for the organised disturbance, by subtracting (5) from (8), and remembering (3b). This gives the following equation:

$$\frac{\partial \tilde{u}_i}{\partial t} + \overline{u}_j \frac{\partial \tilde{u}_i}{\partial x_j} + \tilde{u}_j \frac{\partial \overline{u}_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \tilde{p}}{\partial x_i} + \nu \frac{\partial^2 \tilde{u}_i}{\partial x_j \partial x_j} + \frac{\partial \tilde{r}_{ij}}{\partial x_j}, \tag{9}$$

where $\tilde{r}_{ij}$ is the difference between the ensemble average and time average of $u_i' u_j'$, that is it is the modulation in the Reynolds stress tensor, and is given as follows:

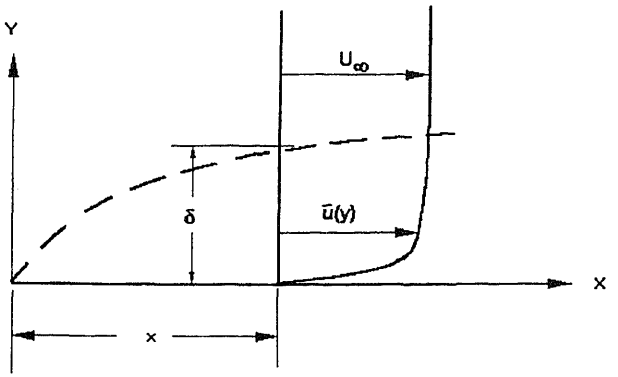


Figure 1. Definition sketch for the flat plate turbulent boundary-layer.

$_j$. It will be instructive to look at simplified physical models and reasoning to understand e manner of resolution proposed for the closure problems.

At this stage we introduce (twice) the rate of strain tensor and (twice) the vorticity tensor, spectively s_{ij} and ω_{ij} , (in generic form) as follows:

$$s_{ij} = \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right); \quad \omega_{ij} = \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right). \tag{11a,b}$$

he expressions in (11) are in generic form, implying that if, for example, u is replaced y $\overline{u}$, then s_{ij} and ω_{ij} are respectively replaced by $\overline{s}_{ij}$ and $\overline{\omega}_{ij}$.

Let us now restrict our attention to 2-D near-parallel flows, specifically the turbulent oundary layer. A definition sketch is shown in figure 1. Note that in the discussions to llow, the vectors (u_1, u_2, u_3) and (u, v, w), will be used interchangably and so also $_1, x_2, x_3)$ and (x, y, z). The x coordinate is in the direction of the free stream, y is the direction normal to the wall, and z is the transverse direction. Also we make the uasi-parallel assumption due to which the mean-velocity field is given as $\overline{u} = \overline{u}(y)$, ≈ 0, and $\overline{w} = 0$. Further, all the mean-velocity gradients, except for $\partial\overline{u}/\partial y$, are either ero, or negligibly small. The turbulence (u_i')-field is also assumed homogeneous in the direction, and near-homogeneous in the x direction. All derivatives of time-averaged uantities are zero in the z-direction, and nearly zero in the x-direction. Moreover the orrelations $\overline{u'w'}$ and $\overline{v'w'}$ are zero. However, $\overline{w'^2}$ is non-zero, and this term keeps the eynolds stress tensor $\overline{u_i'u_j'}$ three dimensional. The velocity scale for the problem could e either the free stream velocity, U_∞ (outer scaling), or the friction velocity, v_* (inner caling). The relevant length scales are the boundary layer thickness δ or the displacement ickness δ^* (both outer scaling), or the viscous length scale, ν/v_* (inner scaling). The haracteristic Reynolds number for the problem is $R \equiv U_\infty\delta/\nu$.

With the problem posed as above, two closure schemes are presented. The first, called he 'isotropic eddy viscosity model' is identical in form to the Newtonian eddy model of eynolds & Hussain (1972). The second, called the 'anisotropic eddy viscosity model' is ased on the model proposed by Pope (1975).

The isotropic eddy viscosity model, for the unperturbed flow, is given by the following well-known expression:

$$-\overline{u'_i u'_j} = -\tfrac{2}{3}k\delta_{ij} + \epsilon\bar{s}_{ij}; \quad \text{with, } k \equiv \tfrac{1}{2}(\overline{u'_i u'_i}). \tag{12a,b}$$

In (12), k is the turbulent kinetic energy, and ϵ is the eddy viscosity. Also, in the wall-bounded two-dimensional flow being considered here, $\epsilon = \epsilon(y)$, that is, the eddy viscosity is a function of the coordinate perpendicular to the wall. Details of the form of ϵ will be discussed later.

Equation (12), despite similarity in form with the corresponding form for laminar shear flow, is no more believed to represent a gradient-type diffusion of momentum. In fact, the justification for the above form is the following (see Tennekes & Lumley 1972). First, the eddy viscosity is of the form $\epsilon \sim v'l$, where l is a turbulent length scale (which may loosely be called the mixing length for historical reasons). This length scale is unique for wall-bounded turbulent shear flows, like in the case of the flat plate, and is given as $l \sim (\kappa y)$ in the inner region and approaching a constant in the outer region. Also κ is the Von Karman constant. Second, the turbulent velocity scale is given uniquely by $v_* = (\tau_w/\rho)^{1/2}$, where τ_w is the wall shear stress; hence v' scales with v_*. Third, the turbulent time scale and the mean flow time scale are of the same order. Thus, form (12) emerges as a dimensional necessity. Since it is not the scope of the present paper to go into a detailed discussion on these points, it suffices to say that form (12) is a valid and legitimate form for wall-bounded shear flows, like the flat plate, that are governed by one velocity scale v_* and one length scale l.

The ideas developed over the past years, on lines of the above discussion, can be summarised by the proposition of the 'generalised eddy viscosity hypothesis' (GEVH). *Such cases of turbulent shear flow that have one characteristic velocity scale and one characteristic length scale of turbulence and, where the turbulence is sustained by the mean shear, will have a time scale of turbulence that is of the same order as the mean flow time scale, the latter being given by the reciprocal of the dominant mean-velocity gradient. Further, the turbulent shear stress tensor in such flows is capable of being described by products of scalar functions derived from the turbulence and appropriate tensorial combinations of the mean-velocity gradients. Further, any description of the shear stress tensor must be tensorially consistent, that is, it must obey the laws of tensor transformation.* In actuality, form (12) is the simplest and most widely used manifestation of GEVH.

What is important for us to note, in the context of the present work, is that if (12) be looked upon as a constitutive equation for $\overline{u'_i u'_j}$, then the right hand side of this equation has one part that is solely dependent on the existing turbulence, viz. $\epsilon \sim v'l$, and the other part on gradients of the mean flow. It is reasonable to assume that if such a flow is perturbed, yet leaving v' and l unchanged, then the eddy viscosity ϵ remains unchanged. In other words, if the existing turbulence is unchanged then ϵ remains unchanged. Further, if the perturbations cause modulations in the mean field, then these modulations appear as modulations of the Reynolds stress tensor. The organised disturbance does exactly this. First it modulates the mean flow, as may be seen from (3b). Second, it can be seen from Reynolds & Hussain (1972) that the modification of the turbulent kinetic energy, due to the

presence of the organised disturbance, is $O(\tilde{u}^2)$. In view of (4) this is negligible, and thus there is no change in the turbulence, and consequently no change in the eddy viscosity.

In the extreme case of a very high-amplitude organised disturbance, possibly even the form (12) would not be valid, since the imposed disturbance would introduce its own length and velocity scales in the problem, thus negating the validity of GEVH.

The consequence of the above discussion is significant, in that (12a) can be extended to the flow modulated by the organised perturbation in the following manner:

$$-\langle u'_i u'_j \rangle = -\tfrac{2}{3}k\delta_{ij} + \epsilon\langle s_{ij}\rangle. \tag{13}$$

It is easy to see now that $\langle u'_i u'_j\rangle$ will be modulated by the organised perturbation *despite* ϵ *having remained unchanged.* It is assumed in (13), following Reynolds & Hussain (1972) that the organised disturbances do not modulate the dilatational part of the stress tensor. Thus $(\langle k\rangle - k)$ is neglected, where $\langle k\rangle = (1/2)\langle u'_i u'_i\rangle$. Besides, even if $(\langle k\rangle - k)$ were to be retained, it would be like the the pressure term in (9), and this would not eventually affect the derivation of the extended Orr–Sommerfeld equation. Further, by subtracting (12a) from (13), it is easy to see that $\tilde{r}_{ij} \sim O(\tilde{u}_i)$, which is consistent with Reynolds & Hussain (1972), and the expression for $\tilde{r}_{ij}$ is given as follows:

$$\tilde{r}_{ij} = \epsilon\tilde{s}_{ij}. \tag{14}$$

Reynolds & Hussain (1972) have not considered $\tilde{u}_i$ as small as in (4) herein, nor have they considered very high amplitude perturbations. However, they have reasoned that, with some restrictions, in their problem also the Newtonian eddy model will give the same expression for $\tilde{r}_{ij}$ as in (14).

2.2 *The anisotropic eddy viscosity model*

The necessity of considering an improved model, over the isotropic eddy viscosity model, is quite strong, and various angles of the matter will be looked into. The main reason for proposing an improved model (without keeping any allied problem, like the hydrodynamic stability problem, in view) has been given by Pope (1975). There could be some flows, wherein the turbulence is isotropic or anisotropic, where GEVH is not valid (one such simple example is given by Tennekes & Lumley (1972), of turbulent flow over a flat-plate with vertical blowing or suction through the plate). Also, there are some flows wherein the turbulence is anisotropic, yet GEVH is valid. The most obvious examples in the latter category are the simple wall turbulent flows, the flat plate boundary layer problem being one of them. The important contribution of Pope (1975) has been to develop advanced level constitutive equations under GEVH, mainly to account for anisotropy in flows that obey GEVH.

First we examine the shortcomings of (12) as a constitutive equation for the Reynolds stress tensor. It can be seen from (12a) that the normal stresses predicted by this model give $\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$ all equal to $\frac{2}{3}k$. This is quite a departure from the known anisotropic behaviour in the entire shear layer and particularly in the wall region. To overcome this shortcoming, the constitutive equation that has been given by Pope (1975), *under the aegis of* GEVH, is as follows:

$$-\overline{u'_i u'_j} = -\tfrac{2}{3}k\delta_{ij} + \epsilon\bar{s}_{ij} - \epsilon C(k/\epsilon_d)[\tfrac{1}{2}[\bar{\omega}_{ik}\bar{s}_{kj} - \bar{s}_{ik}\bar{\omega}_{kj}]]. \tag{15}$$

In (15), ϵ_d is the dissipation rate of the turbulent kinetic energy, and C is a constant to match experimental results. The last term in (15) accentuates the anisotropy generally by augmenting $\overline{u'^2}$, reducing $\overline{v'^2}$, and leaving $\overline{w'^2}$ unchanged. Further, the last term leaves $\overline{u'v'}$ unchanged. Pope (1975) has one more term in (15), involving a scalar invariant of $\bar{s}_{ij}$, namely the scalar product $\bar{s}_{ij}\bar{s}_{ji}$. This term tends to equalise $\overline{u'^2}$ and $\overline{v'^2}$, but makes $\overline{w'^2}$ different from both. This term is rejected for the disturbance equation on two grounds. First, the accentuation of anisotropy in the z direction has not much relevance in the present problem in which basically two-dimensional disturbances are being considered. The second reason is more serious. The $\bar{s}_{ij}\bar{s}_{ji}$-term has to be incorporated artificially with a *two-dimensional* Kronecker delta. Since the latter does not behave like the identity matrix in three dimensions, it destroys the tensorial symmetry of the problem, and, eventually renders the disturbances non-transformable by Squire's theorem. Pope (1975) himself has suggested that this term can be ignored if accentuation of anisotropy in the z direction is not important. Thus, in the present problem, the $\bar{s}_{ij}\bar{s}_{ji}$-term is nominally considered in the (unperturbed) mean-motion only, mainly to estimate the value of the two anisotropy parameters C and C_1, the latter to be introduced later.

Equation (15) has a *unique tensorial form* involving $\bar{s}_{ij}$ and $\bar{\omega}_{ij}$ under GEVH. Over and above the basic isotropic eddy viscosity model, (12), this is the next advanced-level constitutive equation, from amongst the hierarchy of constitutive equations derived by Pope (1975). So long as GEVH is valid, and a model is to be proposed within GEVH, *the uniqueness of the form of* (15) *leaves little scope for seeking out other tensorial forms at the same advanced-level as in* (15). At most, one can look for alternative but equivalent forms for the turbulence-derived *scalar functions* like k/ϵ_d. It is therefore reasonable to conclude, that, since the present problem obeys GEVH, form (15) is clearly both a unique and logical improvement over (12). Also, other such turbulence models that are not within the framework of GEVH, are outside the scope of the present paper. But then, such models would not also admit even a form like (12), and it is doubtful whether such models would capture the real physics in the present problem, as much as GEVH would.

Whilst retaining the basic tensorial form as in (15), for certain conveniences in calculations we introduce an alternative but equivalent scalar function in place of the turbulence-derived scalar function k/ϵ_d in (15). Now, k/ϵ_d represents a turbulent time scale. It is convenient to put this term equivalently as $k/\epsilon_d = \lambda/(\overline{u})'$, where $(\overline{u})' = (\partial\overline{u}/\partial y)$. The function λ is called the anisotropy function, with the constant C in (15) absorbed in λ. The anisotropy function λ thus physically represents the ratio of the turbulent time scale to the mean flow time scale. The constitutive equation adopted in the present problem is therefore given as follows:

$$-\overline{u'_i u'_j} = -\tfrac{2}{3}k\delta_{ij} + \epsilon\bar{s}_{ij} - \epsilon(\lambda/\overline{u}')[\tfrac{1}{2}[\bar{\omega}_{ik}\bar{s}_{kj} - \bar{s}_{ik}\bar{\omega}_{kj}]]. \tag{16}$$

Using exactly the same arguments as in § 2.1 earlier, (16) may be extended to perturbed flow, as follows:

$$-\langle u'_i u'_j\rangle = -\tfrac{2}{3}k\delta_{ij} + \epsilon\langle s_{ij}\rangle - \epsilon(\lambda/\overline{u}')[\tfrac{1}{2}[\langle\omega_{ik}\rangle\langle s_{kj}\rangle - \langle s_{ik}\rangle\langle\omega_{kj}\rangle]]. \tag{17}$$

Also, as before, $\tilde{r}_{ij}$ may be obtained by subtracting (16) from (17), and is given as follows:

A word is now in order regarding the anisotropy function λ, and its efficacy in the present formulation. One can put $\epsilon_d \sim \overline{u'v'}(\overline{u})'$ in (16), the expression for ϵ_d becoming almost exact in the outer region where production nearly equals dissipitation. The expression for λ is thereafter given as follows:

$$\lambda = C(k/(-\overline{u'v'})), \tag{19}$$

where C is either a constant, or at most a slowly varying function. It is noted from (19) that λ is a universal non-dimensional function, and is like the inverse of the structure function given in Townsend (1976), and also plotted by Hinze (1975; figures 7.21 & 7.22). In fact, since k/ϵ_d in (15) is meant to extract a turbulent time scale, one might as well replace ϵ_d by $\overline{u'v'}(\overline{u})'$ in the entire range of y in (15).

The contribution of the $\overline{s}_{ij}\overline{s}_{ji}$-term, to the (unperturbed) mean-motion, is nominally added to (16) using a form similar to (19) with C_1 in place of C. Thereafter, upon expanding (16), one obtains the following expressions for the turbulence stresses:

$$-\frac{\overline{u'^2}}{k} = -\frac{2}{3} + 0 - (C + C_1)\left(\frac{\epsilon(\overline{u})'}{-\overline{u'v'}}\right); \tag{20a}$$

$$-\frac{\overline{v'^2}}{k} = -\frac{2}{3} + 0 + (C - C_1)\left(\frac{\epsilon(\overline{u})'}{-\overline{u'v'}}\right); \tag{20b}$$

$$-\frac{\overline{w'^2}}{k} = -\frac{2}{3} + 0 + 2C_1\left(\frac{\epsilon(\overline{u})'}{-\overline{u'v'}}\right); \tag{20c}$$

$$-\frac{\overline{u'v'}}{k} = 0 + \frac{\epsilon(\overline{u})'}{k} + 0. \tag{20d}$$

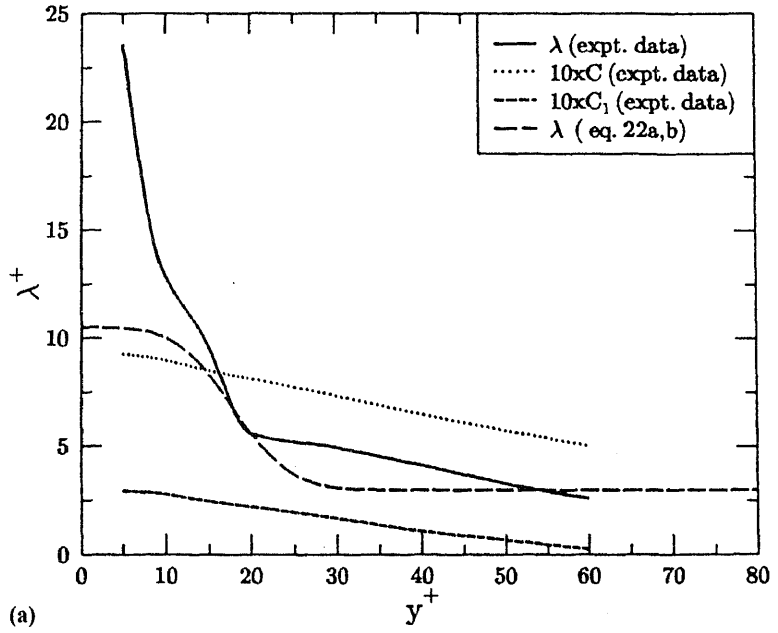
The accentuation of anisotropy now becomes clear, and in view of (20d), (20a,b,c) may be rewritten as follows:

$$\frac{\overline{u'^2}}{k} = \frac{2}{3} + C + C_1; \quad \frac{\overline{v'^2}}{k} = \frac{2}{3} - C + C_1; \quad \frac{\overline{w'^2}}{k} = \frac{2}{3} - 2C_1. \tag{21a,b,c}$$

Estimation of the values of C and C_1 have been carried out from Klebanoff's experimental data (cf. Hinze 1975), and from the DNS data of Spalart (1988). Thereafter, using (19), λ has been evaluated. The results are shown in figure 2a, in which $y^+ \equiv yv_*/\nu$ is y-scaled in inner variables. It is seen that C_1 is quite small except very close to the wall, where it has a value of $\simeq 0.2$. In case that $C_1 = 0$, then it is seen from (21a, b, c) that there is a restriction on admissible values of C, viz. $C \leq 0.67$, to avoid negative values of $\overline{v'^2}$. However, close to the wall, where $C_1 \simeq 0.2$, C can have values $\simeq 0.9$, as may be seen in figure 2a.

It is seen in figure 2a that λ has a value of 25 close to the wall, relaxing to around 2 in the region $y^+ \sim 60$. In the outer region it relaxes further to around 1, and drops to 0 close to the edge of the boundary layer where there is isotropy in the sense that $\overline{u'^2}$, $\overline{v'^2}$, and $\overline{w'^2}$ are nearly equal to each other. The relatively high values of λ obtained close to the wall are because of the strong anisotropy in the wall region.

Stability calculations were performed with a wide latitude of variation in the form of λ, retaining the feature that λ has a high value in the wall region. It was observed that



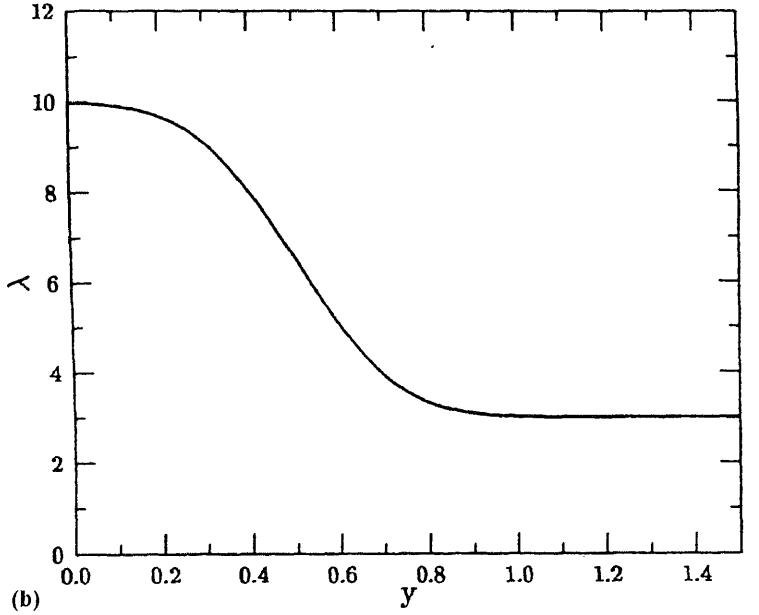


Figure 2. (**a**) Graph of anisotropy function $\lambda^+(=\lambda)$ versus y^+, standardised in terms of inner variables. Based on experimental data of Klebanoff, and from (22). Also shown are C and C_1. (**b**) Graph of anisotropy function λ versus $y(=y_d/\delta)$, standardised in terms of outer variables, based on (23).

the exact shape of the λ-curve is not important. Rather, instability is obtained with a high value of λ ($\simeq 9.0$ or more) in the wall region. Thus instability is due to the anisotropy in the wall region, a feature that could not have been captured by the isotropic eddy viscosity model.

Two of the forms used for λ in the calculations are given below. The first is based on inner variables scaling and is given as follows:

$$\lambda(y^+) = \lambda_{hi}\left[\lambda_{lo}/\lambda_{hi} + \left(\frac{1-\lambda_{lo}/\lambda_{hi}}{2}\right)\left\{1 - \mathrm{erf}\left(\frac{y^+ - y^+_{0\lambda}}{\sigma_\lambda}\right)\right\}\right]; \tag{22a}$$

$$\text{with,} \qquad \lambda_{hi} = 10.5; \quad \lambda_{lo} = 3.0; \quad y^+_{0\lambda} = 18.0; \quad \sigma_\lambda = 7.5. \tag{22b}$$

The λ-curve based on (22) is shown in figure 2a. Equation (22) renders the stability equations stiff and the calculations are more laborious and constrained. This is because the region of high λ is squashed close to the wall. For facility in calculation another expression for λ was used based on outer variable scaling (plot shown in figure 2b) which gives a value of 10 at the wall relaxing much more gently to a value of 3 in the outer region. This equation is given as follows:

$$\lambda(y) = \lambda_{hi}\left[\lambda_{lo}/\lambda_{hi} + \left(\frac{1-\lambda_{lo}/\lambda_{hi}}{2}\right)\left\{1 - \mathrm{erf}\left(\frac{y - y_{0\lambda}}{\sigma_\lambda}\right)\right\}\right]; \tag{23a}$$

$$\text{with,} \qquad \lambda_{hi} = 10.0; \quad \lambda_{lo} = 3.0; \quad y_{0\lambda} = 0.5; \quad \sigma_\lambda = 0.25. \tag{23b}$$

Other forms were also tried, varying the parameters in (23). The fact that instability is not dependent on one particular critical choice of the shape of the λ-curve, but is generally governed by high values of λ in the wall region, is a point that greatly favours the statement of a general proposition: that instability is due to the anisotropy in the wall region.

2.4 *The disturbance equations*

We simplify the evolution equation (9) for $\tilde{u}_i$, remembering that the disturbances are two-dimensional and Squire-transformable, and that the mean-flow is amenable to the quasi-parallel approximation. A stream function ψ is introduced for the organised disturbances such that $\tilde{u} = \partial\psi/\partial y$ and $\tilde{v} = -\partial\psi/\partial x$. After assuming normal modes, ψ may be expressed in the following form:

$$\psi = \phi(y)\exp[i\alpha(x - ct)], \tag{24}$$

where, α is the spatial wavenumber and $c = c_r + ic_i$ is the (complex) wave speed. Introducing (14) in the evolution equation (9) for $\tilde{u}_i$, remembering the quasi-parallel approximation for the mean flow, and remembering the closure equations (14, 18) for $\tilde{r}_{ij}$, one arrives at extended forms of the Orr–Sommerfeld equation (a more detailed derivation may be found in Sen & Veeravalli 1994). For the isotropic eddy viscosity model this is given as follows:

$$i\alpha[(\overline{u} - c)(\phi'' - \alpha^2\phi) - \overline{u}''\phi] - (1/R)[\phi'''' - 2\alpha^2\phi'' + \alpha^4\phi]$$
$$-(1/R)[E\{\phi'''' - 2\alpha^2\phi'' + \alpha^4\phi\} + 2E'\{\phi''' - \alpha^2\phi'\}$$

follows:

$$
\begin{aligned}
&i\alpha[(\overline{u}-c)(\phi''-\alpha^2\phi)-\overline{u}''\phi]-(1/R)[\phi''''-2\alpha^2\phi''+\alpha^4\phi]\\
&-1/R[E\{\phi''''-2\alpha^2\phi''+\alpha^4\phi\}+2E'\{\phi'''-\alpha^2\phi'\}\\
&+E''\{\phi''+\alpha^2\phi\}]+(\lambda E/R)[-2i\alpha\phi'''+2i\alpha^3\phi']\\
&+(2i\alpha\phi'/R)[\lambda E''+2\lambda' E'+\lambda'' E]=0. \qquad (26)
\end{aligned}
$$

In both (25) and (26), primes ($'$) denote differentiation with respect to y. All quantities in (25) and (26) have been non-dimensionalised by U_∞, δ and ν, i.e. $E \equiv \epsilon/\nu, \overline{u} \equiv \overline{u}_d/U_\infty, y \equiv y_d/\delta$ etc., where subscript 'd' represents dimensional or physical quantities. Also, in (25) and (26), the first group of terms in square brackets corresponds to the Rayleigh equation; the first two groups of terms in square brackets correspond to the classical Orr–Sommerfeld equation; and the remaining terms constitute the modifications in the classical Orr–Sommerfeld equation. Now, $E', E'' \to 0$ (to be shown later) at the edge of the boundary layer, and, λ can be made to obey the less stringent conditions $\lambda', \lambda'' \to 0$ (without letting $\lambda \to 0$) at the edge of the boundary layer. Thus, the appropriate boundary conditions for the two extended Orr–Sommerfeld equations, viz. (25) and (26), are the same as those for the classical Orr–Sommerfeld equation, namely:

$$\phi, \phi' = 0 \text{ at } y = 0; \qquad (27a)$$

$$\phi(y) \sim e^{-\alpha y}, \text{ for } y \to \infty. \qquad (27b)$$

Either of the disturbance equations (25) or (26), along with the appropriate boundary conditions (27), constitutes an eigenvalue problem. In the temporal problem, α and R are chosen real and $c = c_r + ic_i$ is determined as the (complex) eigenvalue. Stability or instability is obtained respectively as $c_i < 0$ or $c_i > 0$. As noted earlier the extended Orr–Sommerfeld equation, as given by (25), did not yield any unstable modes although near neutral modes were observed.

Next we look at the inner-variable form of (26). It is well known that at high values of R, the turbulence in the region close to the wall assumes a universal character, which is best seen by expressing the governing equation in inner variables. The velocity and length scales are v_* and ν/v_*, respectively. Denoting quantities scaled by inner variables by the superscript (+), we note that, $\lambda = \lambda^+; E = E^+; B \equiv v_*^2/U_\infty^2; y^+ = yR\sqrt{B}; \overline{u}^+ = \overline{u}/v_*$; and, $\alpha^+ = \alpha/(R\sqrt{B})$, where B is the non-dimensional wall shear stress. Thus, in inner variables, the disturbance equation (26) becomes:

$$
\begin{aligned}
&i\alpha^+[(\overline{u}^+-c^+)(\phi''-\alpha^{+2}\phi)-\overline{u}^{+\prime\prime}\phi]-[\phi''''-2\alpha^{+2}\phi''+\alpha^{+4}\phi]\\
&-[E^+\{\phi''''-2\alpha^{+2}\phi''+\alpha^{+4}\phi\}+2E^{+\prime}\{\phi'''-\alpha^{+2}\phi'\}\\
&+E^{+\prime\prime}\{\phi''+\alpha^{+2}\phi\}]+\lambda^+E^+[-2i\alpha^+\phi'''+2i\alpha^{+3}\phi']\\
&+2i\alpha^+\phi'[\lambda^+E^{+\prime\prime}+2\lambda^{+\prime}E^{+\prime}+\lambda^{+\prime\prime}E^+]=0; \qquad (28)
\end{aligned}
$$

where, in (28), primes ($'$) denote differentiation with respect to y^+. Note that (28) does not show any explicit dependence on the Reynolds number, R. As will be seen later, (28)

corresponding equation (28) would indicate that the behaviour of these wall modes should be near universal at high R. This is because the outer region does not have a strong influence on wall modes. Universality is actually confirmed by the present numerical results.

3. Details of various functions

The mean flow in wall turbulence is known to be multi-layered. Closest to the wall is the viscous sub-layer, followed by the buffer zone, and then followed respectively by the inner and outer log-law regions. Mathematical expressions are available for the mean velocity distributions, individually for each of these respective zones; but taken together these expressions do not constitute an analytically continuous expression for the mean velocity distribution. For hydrodynamic stability calculations as in the present problem, it is necessary to obtain an analytically continuous expression for the mean flow.

One very successful method of generating an analytically continuous mean velocity distribution for turbulent channel flow is reported in the work of Reynolds & Tiederman (1967). In the present work, the mean velocity profile for turbulent boundary layer flow is generated on lines of the work of Reynolds & Tiederman (1967). This requires obtaining suitable expressions for the eddy viscosity, and the shear stress distribution. This is discussed next.

3.1 *The expression for eddy-viscosity*

First we look at the eddy viscosity $E(y)$. The expression for the eddy viscosity $E(y)$ is given by a combined expression of Van Driest's law for the wall region and Reichardt's expression for the outer region. The composite $E(y)$ is given in Reynolds & Tiederman (1967), and Reynolds & Hussain (1972). Also, a slightly different expression is given in Hussain & Reynolds (1975). The Reynolds & Tiederman (1967) expression is given as follows:

$$E(y) = \frac{1}{2}\left[1 + \kappa^2 \frac{R^2 B}{9}(2y - y^2)^2(3 - 4y + 2y^2)^2 \times \left\{1 - \exp\left(\frac{-yR\sqrt{B}}{A^+}\right)\right\}^2\right]^{\frac{1}{2}} - \frac{1}{2}. \tag{29}$$

In (29), κ is the Von Karman constant, A^+ is the Van Driest constant. The above equation is best suited for channel flow and a detailed plot of $E(y)$ may be seen in Reynolds & Tiederman (1967). A slightly modified basic expression for $E(y)$ is needed for boundary layer flow, because E', $E'' \simeq 0$ for $y = 1$ (boundary layer edge). This basic expression, denoted by E_b, again needs to be further modified to account for the intermittency in the

outer region. But first we define the basic expre ion E_b, as follows:

$$E_b(y) = \frac{1}{2}\left[1 + \kappa^2 \frac{R^2 B}{4}(2y - y^2)^2(2 - 2y + y^2)^2 \times \left\{1 - \exp\left(\frac{-yR\sqrt{B}}{A^+}\right)\right\}^2\right]^{\frac{1}{2}} - \frac{1}{2}. \tag{30}$$

In the expression for E_b, i.e. in (30), κ has been chosen as $\kappa = 0.4$, and A^+ has been chosen as $A^+ = 27$. Next we see in (30), what further modifications in E_b are needed in order to accommodate the intermittency in the outer region. Although E_b satisfies E_b', $E_b'' \simeq 0$, at $y = 1$, it still does not approach the right numerical value of E at $y = 1$. So, the first correction is to multiply E_b by an error function $g(y)$, so that the right value of the eddy viscosity is obtained at $y = 1$, prior to correction for intermittency. Next, to account for external intermittency of the boundary layer, for $y > 0.4$, the eddy viscosity expression is further multiplied by the intermittency function (another error function) $\gamma(y)$. The final expression for eddy viscosity is given as follows:

$$E(y) = E_b(y)g(y)\gamma(y); \tag{31a}$$

where,

$$g(y) = (1 - 2a) + a\left[1 - \text{erf}\left(\frac{y - y_{0g}}{\sigma_g}\right)\right];$$
$$\text{with,} \quad a = 0.2; \quad y_{0g} = 0.4; \quad \sigma_g = 0.13; \tag{31b}$$

and,

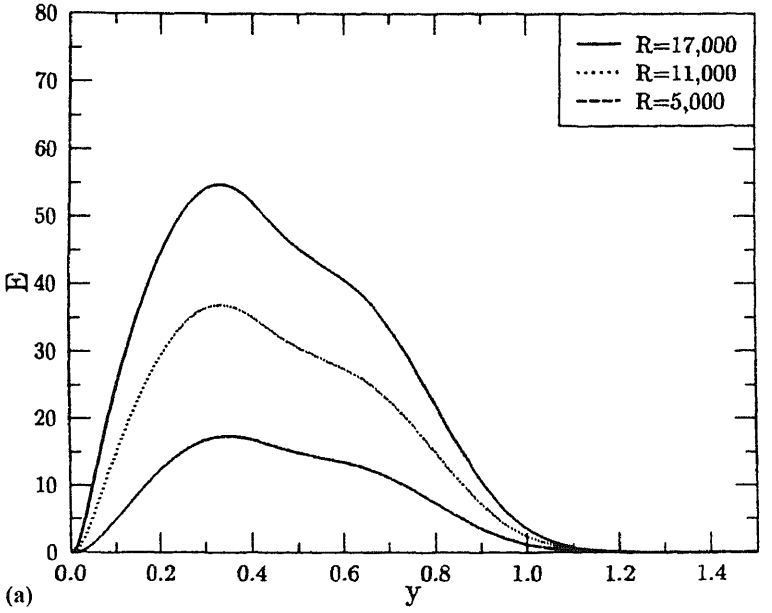
$$\gamma(y) = 0.5\left[1 - \text{erf}\left(\frac{y - y_{0\gamma}}{\sigma_\gamma}\right)\right];$$
$$\text{with,} \quad y_{0\gamma} = 0.797; \quad \sigma_\gamma = 0.204. \tag{31c}$$

Plots of $E(y)$, for different values of R, and corresponding to (31), are shown in figure 3a. Also, figure 3b shows the plot for $\epsilon/(v_*\delta)$ for $R = 75,000$, and, shown in this figure also are the points corresponding to the measurements by Klebanoff (1954), reported in Hinze (1975).

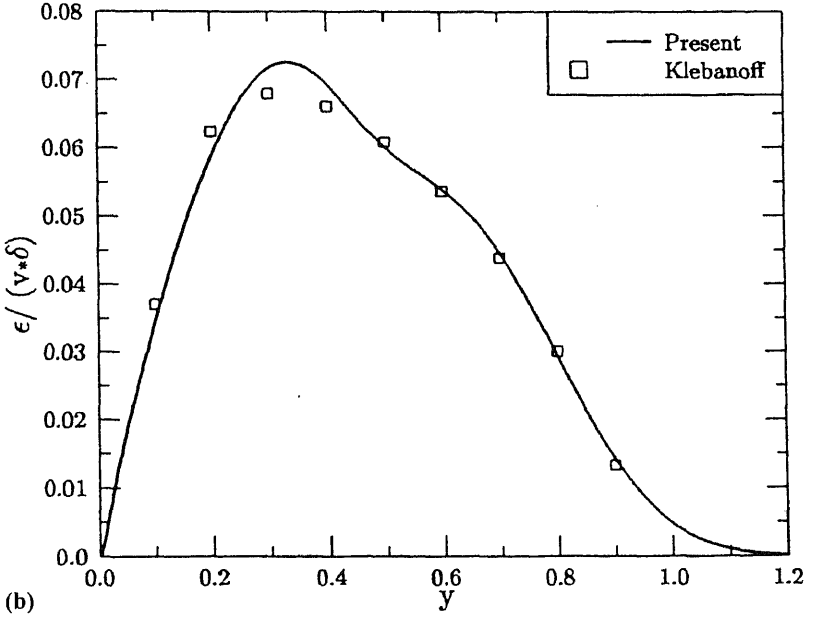
One may look at the form of E_b in inner variables, to make an assessment of the (marginal) departure from universality, of the disturbance equation (26) and (28) because of the outer scaling. Thus using inner scaling in (30) we have

$$E_b{}^+ = \frac{1}{2}\left[1 + \kappa^2 \frac{y^{+2}}{4}(2 - y)^2(2 - 2y + y^2)^2 \left\{1 - \exp\left(\frac{-y^+}{A^+}\right)\right\}^2\right]^{\frac{1}{2}} - \frac{1}{2}. \tag{32}$$

Note in (32) that $y \equiv y^+/(R\sqrt{B})$. Therefore, the outer region does contribute to E_b^+, and also therefore, E^+ is dependent on the Reynolds number R. Nevertheless, as stated



(a)



(b)

Figure 3. (Continued.)

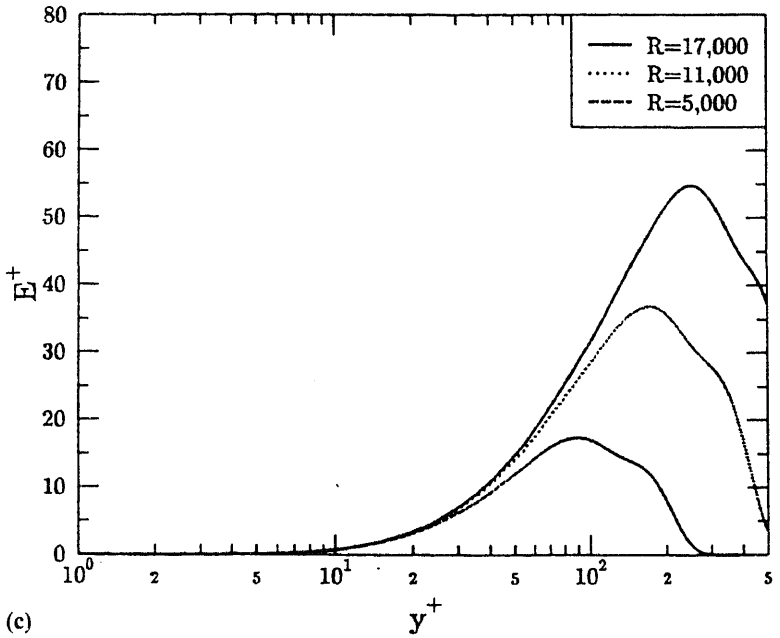


Figure 3. (**a**) Graph of eddy viscosity $E(=\epsilon/\nu)$ versus $y(=y_d/\delta)$. (**b**) Graph of eddy viscosity $\epsilon/(\upsilon * \delta)$ versus $y(=y_d/\delta)$. Comparison with Klebanoff's data at Re = 75,000. (**c**) Graph of eddy viscosity $E^+(=E)$ versus y^+, inner variables. Scale of y^+ is logarithmic.

wall mode that is obtained in the problem. One may now similarly express (31) in inner variables scaling, to obtain $E^+(y^+)$ (which is the same as $E(y^+)$). Figure 3c shows plots of the eddy viscosity $E^+(y^+)$, in inner variables. This figure shows that in the *inner region* the eddy viscosity is the same across different R, a feature that is responsible for the universality of the wall modes.

3.2 *The expression for the shear stress distribution*

The distribution for the (total) shear stress τ is needed in order to be able to integrate the equation of motion to obtain the mean velocity distribution $\overline{u}(y)$. It is well known from experimental data that the distribution of mean shear looks similar to a tanh-profile, or a complementary error function profile. The expression chosen here is again a composite expression comprising an error function and a Gaussian function, to be able to match the experimental shear stress distribution, and, also, for the mean velocity profile to be able to satisfy the proper outer boundary conditions. This composite expression is given as follows:

$$\tau(y) = \tau_w F(y) G(y); \tag{33a}$$

where, τ_w is the wall shear stress. Also τ is dimensional or non-dimensional depending on whether τ_w is dimensional or non-dimensional. When τ_w is non-dimensional then $\tau_w = B$. Further,

$$F(y) = 1 + b\left[1 + \text{erf}\left(\frac{y - y_{0F}}{\sigma_F}\right)\right];$$
$$\text{with,} \qquad b = 0.7; \quad y_{0F} = 0.56; \quad \sigma_F = 0.204; \tag{33b}$$

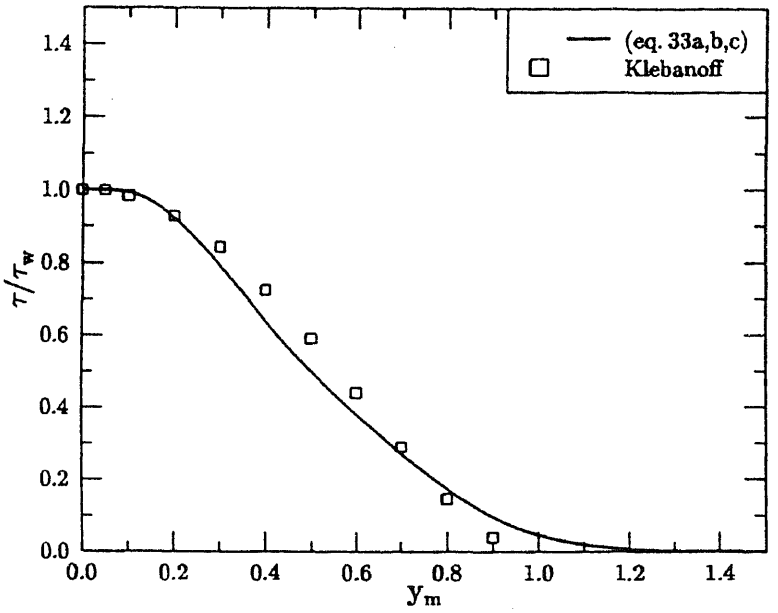


Figure 4. Graph of shear-stress τ/τ_w versus y_m. Here y is rescaled so that at $y_m = 1$, $\overline{u} = 0.99$. Data points of Klebanoff are also shown, for Re = 75,000.

and,

$$G(y) = 1,\ G'(y) = 0; \quad \text{for } 0 \le y < y_{0G};$$

$$G(y) = \exp\left[-\frac{1}{2}\left(\frac{y - y_{0G}}{\sigma_G}\right)^2\right]; \quad \text{for } y \ge y_{0G};$$

$$\text{with,} \qquad y_{0G} = 0.05; \quad \sigma_G = 0.26. \tag{33c}$$

A typical plot of τ/τ_w versus y is shown in figure 4. In this figure the scale for y is modified to y_m (defined in the next sub-section). Also shown are data points for τ/τ_w measured by Klebanoff (1954), and as reported in Hinze (1975).

3.3 *Generation of the mean velocity profile*

The x-momentum equation for mean velocity simplifies for boundary layer flow to the following:

$$\overline{u}\frac{\partial \overline{u}}{\partial x} + \overline{v}\frac{\partial \overline{u}}{\partial y} = \frac{1}{R}\frac{\partial}{\partial y}\left[(1+E)\frac{\partial \overline{u}}{\partial y}\right] = \frac{\partial \tau}{\partial y}, \tag{34}$$

In the discussion that follows, the convective acceleration terms are not required. Hence (34) gives a relationship between E and τ (which is in non-dimensional form). Thus, after integrating (at any given station in x), (34) becomes the following:

$$(1+E)\frac{\partial \overline{u}}{\partial y} = R\tau; \tag{35}$$

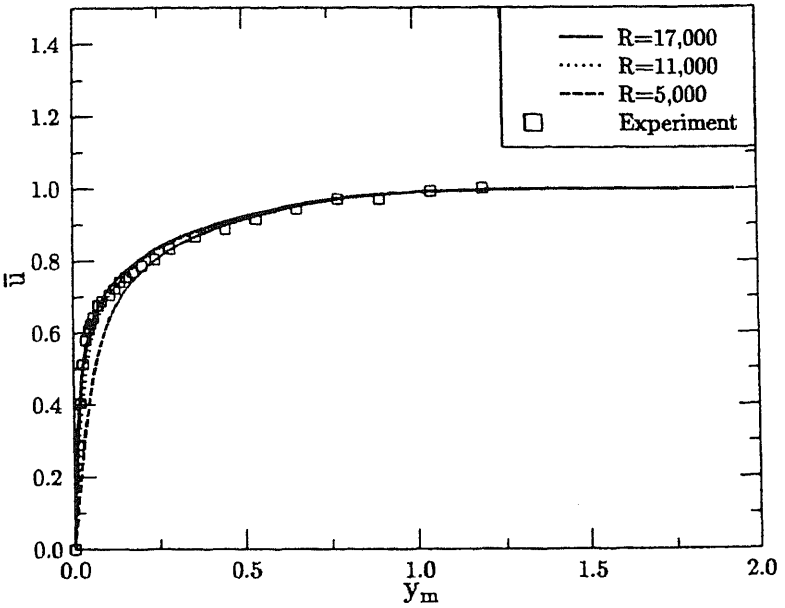


Figure 5. Graph of mean-velocity u versus y_m. y rescaled so that at $y_m = 1$, $\overline{u} = 0.99$. Data points are from present experiments, for Re around 17,000.

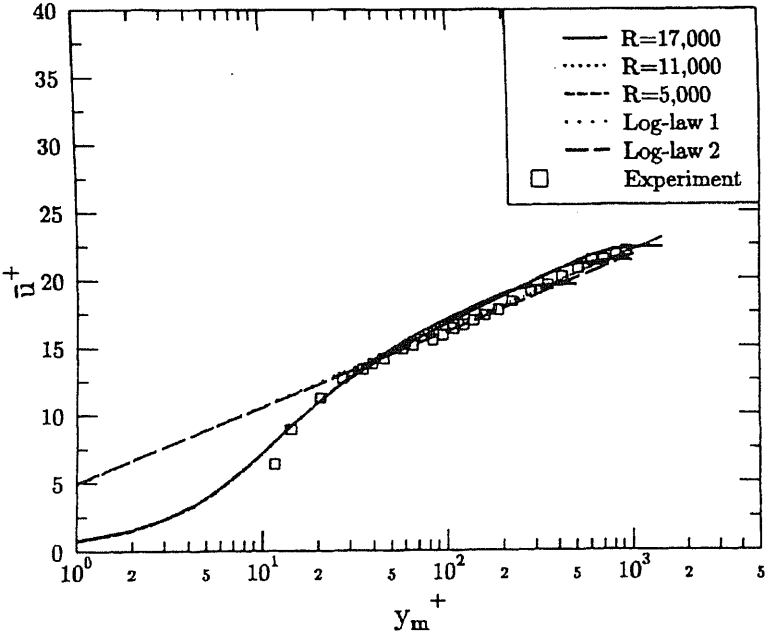


Figure 6. Graph of $\overline{u}^+$ versus y_m^+. Also $y_m^+ = y_m u*/\nu$, where y is rescaled so that at $y_m = 1$, $\overline{u} = 0.99$. Log-law $1 : \overline{u}^+ = 5.0 + 5.75\log(y^+)$. Log-law $2 : \overline{u}^+ = 5.0 + 5.62\log(y^+)$.

where, in (35), τ is given from (33a,b,c). Also, it may be noted in (35) that $\partial\overline{u}/\partial y$ satisfies the condition $\partial\overline{u}/\partial y = RB$, for $y = 0$. Also $\partial\overline{u}/\partial y$ satisfies the condition $\partial\overline{u}/\partial y \to 0$ for $y \to \infty$; because, $E \to 0$ for $y \to \infty$, and, the right hand side of (35), viz. $R\tau \to 0$ for $y \to \infty$. Integrating (35) we obtain

$$\overline{u} = RB \int_0^y \frac{F(y')G(y')}{1 + E(y')}\,\mathrm{d}y'; \tag{36}$$

where, in (36), $F(y)$, $G(y)$ are given from (33a,b,c). Since $\overline{u}$ must approach the free stream velocity $\overline{u}_\infty$ (= 1 in non-dimensional form), therefore, it follows from (36) that,

$$1 = RB \int_0^\infty \frac{F(y)G(y)}{1 + E(y)}\,\mathrm{d}y. \tag{37}$$

Thus iterating between (30), (31a,b,c), (33a,b,c), and (37) one obtains the converged value of B. After this the expression for $\overline{u}$ is given from equation (36).

Typical plots of the mean velocity distribution $\overline{u}$, both in outer and inner variables are shown respectively in figures 5 and 6. Also shown in these figures are the points corresponding to the present experimental measurements. Solution for $\overline{u}$ from (36) is given in terms of $y(\equiv y_\mathrm{d}/\delta)$. Thus $y = 1$ represents $y_\mathrm{d} = \delta$. Now, after $\overline{u}$ has been calculated from (36), it is not guaranteed that at $y = 1$, $\overline{u}$ has the exact value of $\overline{u} = 0.99$. Therefore, for comparison with experiments, it is sometimes expedient to rescale y to y_m, where $y_m \equiv y/d_{99}$, and, for $y = d_{99}$, $\overline{u} = 0.99$. The scale for y, in figures 5 and 6, is thus modified to y_m, for ease of comparison with experimental results. Since the Reynolds numbers (in figures 5 and 6) are relatively small, the inner and outer regions are not well separated. Therefore, one need not expect the log-law to be valid over an extended region in y_m^+. This may also be confirmed from the DNS data of Spalart (see figure 5 of Spalart 1988). In figure 6 it may be noted that there is very good agreement in the inner-wall region. This is what is really important in the stability calculations because the critical point lies in the inner-wall region. In the outer region also, there is good agreement with the experimental data points.

It is clearly seen from the above figures that a very satisfactory method of generating an analytically continuous mean velocity has been obtained. Further refinements, if needed, may be made by adjusting the constants in the expressions for E and τ.

4. Numerical results

The numerical solutions of the extended Orr–Sommerfeld equation (26) were based on finite-difference methods developed in the earlier works of one of the authors (see Sen & Vashist (1989) and Sen 1993). In the present problem, the methods described in the earlier works for the stability of the Blasius boundary layer are extended to the present turbulent boundary layer problem. The resolution used is ten times greater, and the basic step size in y, in the present problem, is $h = 0.001$. Over a range of $0 \leq y \leq 1.5$, 1501 points are required for each array specification. A seven-point finite difference scheme, employing a Noumerov transform and using a molecule described in Sen *et al* (1985), was employed in the calculations, and this gives basic round-off accuracy as $O(h^6)$, i.e.

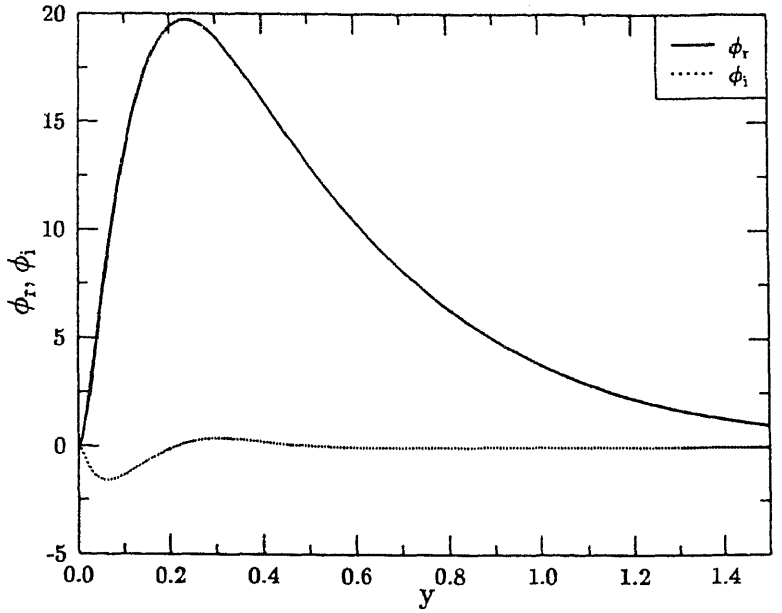


Figure 7. Graph of eigenfunction $\phi = \phi_r + i\phi_i$ versus y. Normalisation $\phi = 1 + 0i$ at $y = 1.5$. Also Re = 5000, $\alpha = 2.7$, $c_r = 0.3486034$, $c_i = 0.0044436$.

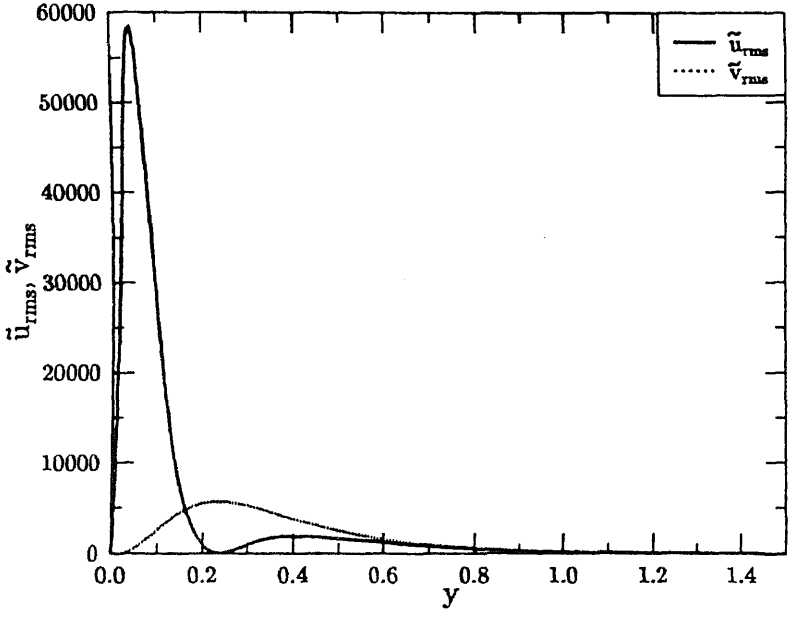


Figure 8. Graph of r.m.s. values of $\tilde{u}$, $\tilde{v}$ versus y. Normalisation, $\phi = 1 + 0i$ at $y = 1.5$. Also, Re = 5000, $\alpha = 2.7$, $c_r = 0.3486034$, $c_i = 0.0044436$.

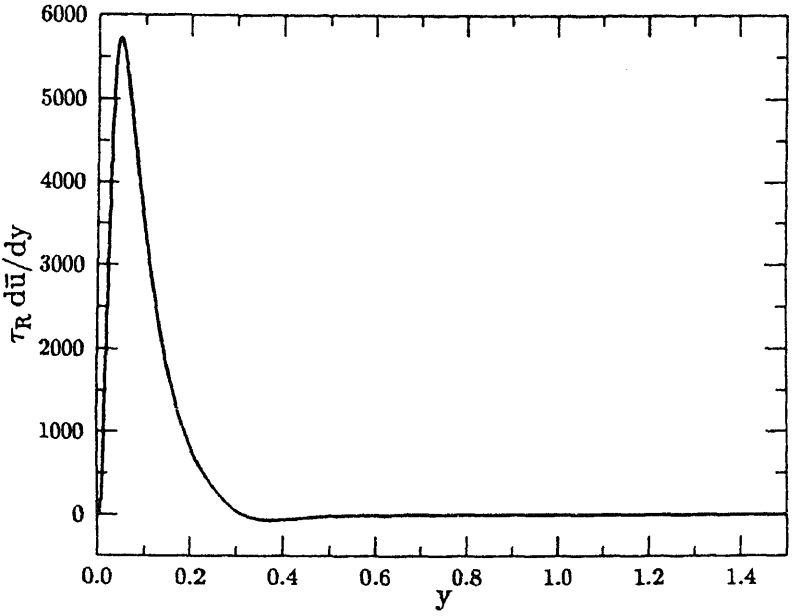


Figure 9. Graph of $\tau_R\, d\bar{u}/dy$ versus y. Normalisation, $\phi = 1 + 0i$ at $y = 1.5$. Also, Re = 5000, $\alpha = 2.7$, $c_r = 0.3486034$, $c_i = 0.0044436$.

10^{-18} with $h = 0.001$. Thus all care was taken to ensure that the very rapid changes taking place in the inner region, are faithfully taken into account in the numerical calculations. Also, the Blasius boundary layer results were verified as a cross-check. Calculations were performed using double precision arithmetic.

The numerical solution of the extended Orr–Sommerfeld equation (26) was obtained for a wide range of wavenumbers α, for R equal to 5000, 11000 and 17000. At larger values of R the sizes of the arrays become prohibitively large, if good resolution is desired close to the wall. However, from (28), it is clear that if one is interested in wall modes it is not really necessary to compute at very high R. Indeed, it is seen from the results presented below that trends become nearly universal in the reported range of R.

For all the three Reynolds numbers noted above, unstable eigenvalues were found for a wide range of α. A typical set of results is shown in figures 7 and 8, for the point defined by (2.7, 5000) in the (α, R) plane. Here, $c_r = 0.349$ and $c_i = 0.0044$. The value of c_r indicates that this is a wall mode. (Note: $\overline{u} = 0.35$ for $y = 0.03$, when $R = 5000$.) Figure 7 shows the eigenfunction ϕ, normalised as $\phi = 1 + 0i$ at $y = 1.5$. Figure 8 shows plots of the r.m.s. values of $\tilde{u}$ and $\tilde{v}$ as estimated from the numerical solution. As seen from figure 8, r.m.s. values of $\tilde{u}$ peaks at $y^+ \approx 11$, which agrees extremely well with experimental data (see Tennekes & Lumley 1972) and the results of direct numerical simulations (Spalart 1988). Similar results were obtained for the production term, $\tau_R(\overline{u})'$, where $\tau_R = \overline{\tilde{u}\tilde{v}}$, which peaks at $y^+ \approx 12.5$, as shown in figure 9. (The inordinately high values of ordinates in figures 8 and 9 are no cause for concern, and follow from the normalisation $\phi = 1 + 0i$ at $y = 1.5$.)

Figure 10 shows the growth rate of the organised disturbance, $\alpha^+ c_i^+$, plotted against the wavenumber α^+. One sees that the band of unstable wavenumbers is fairly large and

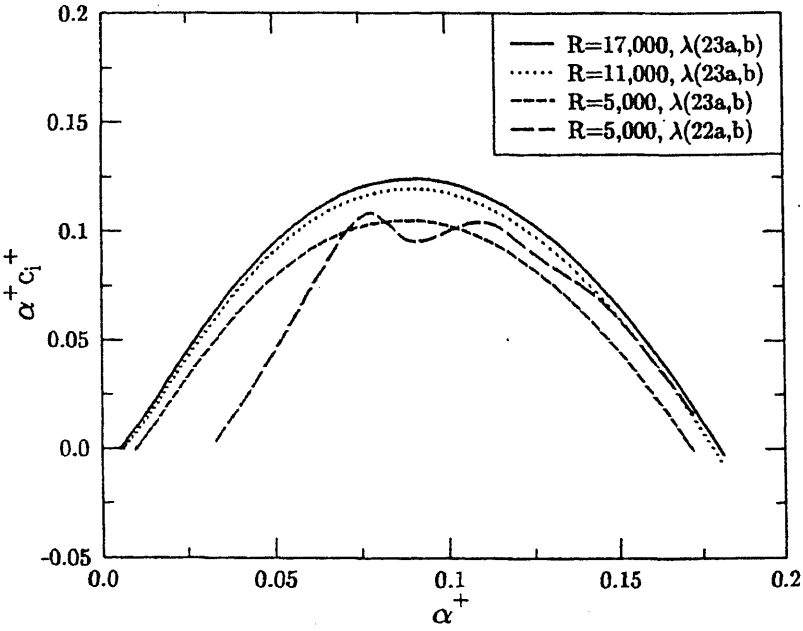


Figure 10. Graph of the growth rate, $\alpha^+ c_i^+$ versus α^+, in inner variables scaling.

virtually the same, in inner variables, for all the three values of R. Secondly, in keeping with (26), the growth rate curves seem to reach, asymptotically, a limiting curve. Most of the calculations reported are with λ based on (23). However, one growth rate curve is also shown in the figure, with λ based on (22). As mentioned earlier, calculations using the latter expression of λ are stiffer and therefore not persisted with for higher R. Nevertheless, the growth rate curves show that the trend of the results, and the range of unstable α, are very comparable by these two widely different shapes of the λ-curves. Thus two important conclusions drawn are as follows. First, the instability is governed by a large value of λ ($\simeq 9$) in the inner region. Secondly, the exact shape of the λ-curve is not important.

5. Experiments

The experiments were conducted in the 0.61 m × 0.61 m × 4.75 m test section, suction tunnel of the Gas Dynamics Laboratory of this department. Figure 11 shows a schematic drawing of the tunnel and the coordinate system used. The boundary-layer growing along the bottom wall of the test section was chosen for measurement. A boundary-layer trip (1.5 mm square rod) was placed at the start of the test section. Other details on the tunnel and the instruments used for measurements may be found in Sen *et al* (1993)and Sen & Veeravalli (1996a).

Measurements were made at a nominal tunnel speed of approximately 8 m/s and at an x

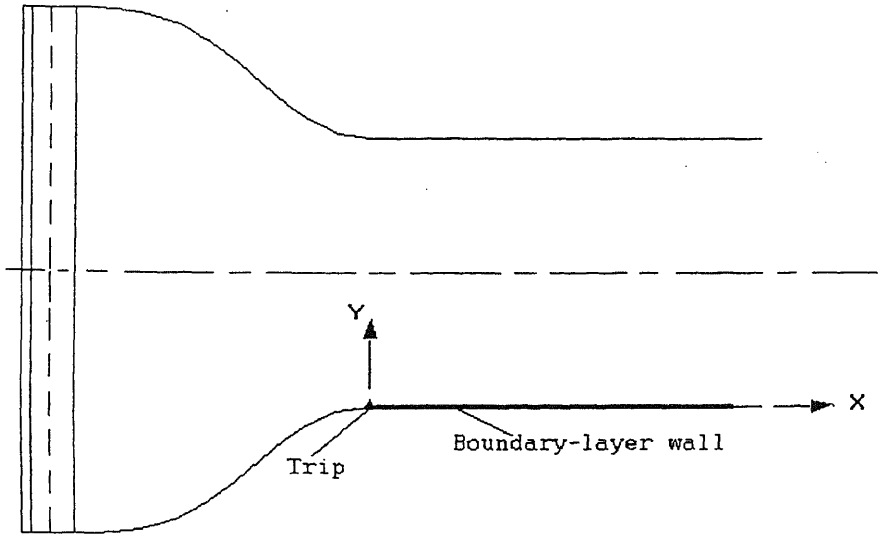


Figure 11. Schematic diagram of the experimental set-up.

the wall. The mean velocity profiles used in the numerical work and the best fit log-layer straight line from Purtell *et al* (1981) are also shown in figure 6. As mentioned before, since R is not very high, the inner log-law region does not extend over a very wide range in y^+.

The friction velocity v_* in the present experiments was estimated by several different methods. The best-fit log-layer straight line to various boundary layer experiments at comparable Reynolds numbers, reported in literature is (Purtell *et al* 1981):

$$\overline{u}^+ = 5.0 + 5.62 \log_{10} y^+. \tag{38}$$

where, $\overline{u}^+ \equiv \overline{u}/v_*$ and $y^+ \equiv y_d v_*/\nu \equiv yR\sqrt{B}$.

The value of v_* for the present experiments may thus be estimated from both the slope and the intercept of the above log-law. From the slope one obtains $v_* = 0.352$ m/s while from the intercept the value for v_* is 0.353 m/s. The friction velocity may also be estimated from the Reynolds number R_θ, using the skin friction data reported in literature (Purtell *et al* 1981). Using this method one obtains a value of 0.344 m/s. Finally, the Preston-tube measurement using the calibration of Patel (1965) yielded $v_* = 0.352$ m/s. We see that

Table 1. Boundary-layer parameters.

$$\delta^* \equiv \int_0^\delta (1 - \overline{u}/U_\infty)\mathrm{d}y = 4.61\,\mathrm{mm}$$

$$\theta \equiv \int_0^\delta \frac{\overline{u}}{U_\infty}(1 - \overline{u}/U_\infty)\mathrm{d}y = 3.28\,\mathrm{mm}$$

the different methods for estimating the friction velocity agree to within 2.5%. For the plots based on experiments in figures 5 and 6, v_* has been taken as 0.352 m/s.

The boundary layer thickness, defined for the experimental results as the location at which the mean velocity reaches $0.99U_\infty$, is 3.35 cm for this x location. Various parameters obtained from the mean velocity profile like the displacement thickness δ^* momentum thickness, θ etc. are reported in table 1. The shape factor, H, falls well within the range reported in Purtell *et al* (1981) and agrees to better than 1% with the data of Coles (1962), reported in Purtell *et al* (1981).

Figure 12 shows compensated, one-dimensional, power spectra of the u' velocity fluctuations, $\kappa^+ E_{11}^+(\kappa^+)$, plotted against κ^+. Here $\kappa \equiv 2\pi f/\overline{u}$, where f is the frequency,

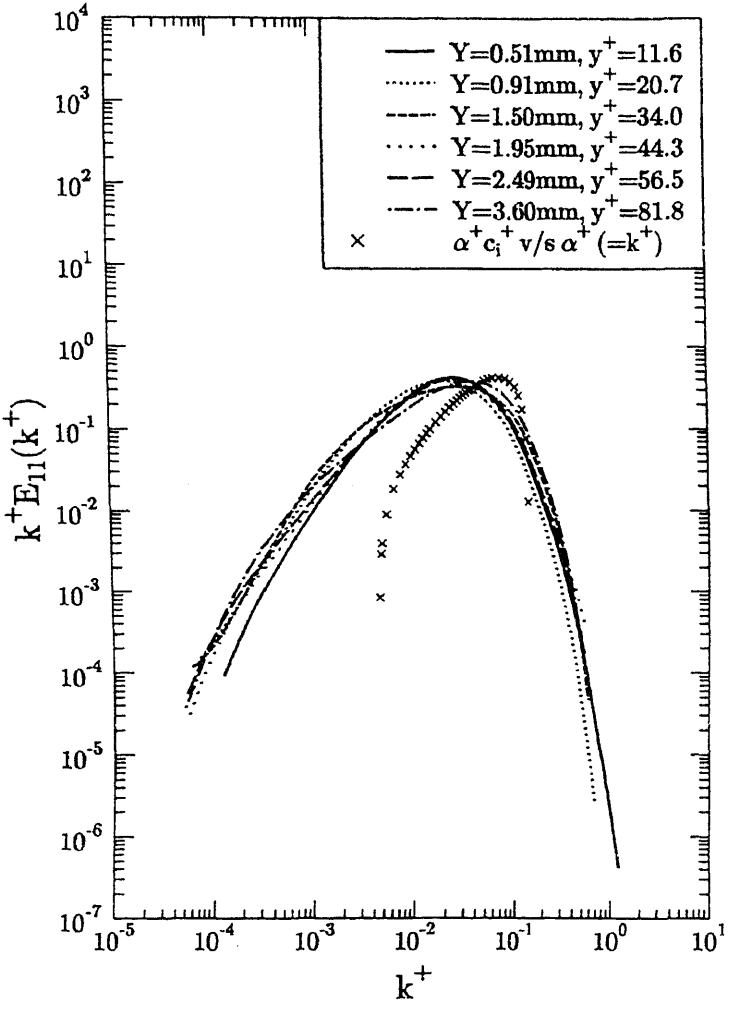


Figure 12. Graph of compensated power spectra $k^+E_{11}(k^+)$ versus k^+, of the u' fluctu-

a single curve, as expected. Also shown in the same figure is the plot of the analytically obtained growth rate curve of $\alpha^+ c_i^+$ versus κ^+. Too much should not be read into the location of the 'peak' of this curve, as this curve only gives the growth *rates* according to linear stability theory. Nevertheless, the $\alpha^+ c_i^+$ versus k^+ curve is contained fully within the collapsed experimental spectral curve, and within the energy-containing eddies region. This is yet another feature that keeps alive the possibility of some connection between the instability waves and actual turbulence.

6. Conclusions

An extended Orr–Sommerfeld equation has been derived to accurately describe the evolution of organised disturbances in the presence of background turbulence. This has been done based on a few novel features in the formulation. First it has been emphasized that the turbulent stresses in the problem can be suitably modelled based on the generalised eddy viscosity hypothesis (GEVH). This is a very significant observation to make, and this is the basis for legitimacy of the formulation of Reynolds & Hussain (1972), as well as of the present formulation. Second, after stipulating the validity of GEVH in the present problem, a *significant and logical* improvement has been made in modelling the Reynolds stress tensor, over Reynolds & Hussain (1972), based on the generalised anisotropic formulation of Pope (1975). This improvement is also *unique* in form. Thirdly, in the formulation, the anisotropy is properly accounted for in terms of a universal non-dimensional function λ, which has been called the 'anisotropy function' herein.

The numerical solution of the extended Orr–Sommerfeld equation (26) yielded unstable eigenvalues over a wide range of α. It has been found that instability is obtained by a large value of $\lambda (\simeq 9$ or more) in the wall region. However, the calculations are not sensitive to the exact shape of the λ-curve. The mode of instability obtained is the wall-mode, and the eigenfunction and the r.m.s. distributions of $\tilde{u}$ and $\tilde{v}$ are found to be qualitatively similar to those obtained for Blasius flow. Moreover, the locations of the respective peaks of the $\tilde{u}$ r.m.s. profile and the production term $\tau_R (\overline{u})'$, agree well with experimental values reported in literature, for *actual turbulence*. Thus the organised disturbances seem to mimic some of the features of actual turbulence.

The band of unstable eigenvalues, when scaled with respect to inner variables, indicates that results approach universality with high R. Particularly, the growth rate curves for different R, in inner variables, viz. $\alpha^+ c_i^+$ versus α^+, approach a limiting curve for high R, as may be seen in figure 10. This is also a feature that correlates well with the inner scaling in actual turbulence.

The experimental rig produces an adequately well-conditioned flat plate turbulent boundary layer flow to permit comparison with and validation of the numerical results. Detailed measurements of the one-dimensional power spectra for u' were taken at different y^+, in the buffer and log regions. These spectra were found to collapse well into a single curve. The limiting growth rate curve of the instabilities, plotted as $\alpha^+ c_i^+$ versus κ^+, is seen to be contained entirely within the power spectral curve of the u' fluctuations, and the

entire range of unstable α^+ is contained within the energy containing eddies range. This is yet another feature that seems to support the possibility of some connection between the instabilities and actual turbulence.

As a final word in conclusion, it may be stated that the present work keeps alive the question of a possible connection between stability theory and actual turbulence, in wall-bounded turbulent flows.

7. Dedication

We wish to dedicate this paper to Professor Peter Bradshaw FRS (somewhat belatedly) on the occasion of his 60th birthday. One of us (SVV) would also like to take this opportunity to thank Professor Bradshaw for all that he has learnt from him, through his books and lectures, and through discussions, on various theoretical and experimental aspects of turbulence, and on turbulent boundary layers in particular.

This project was funded by the Council of Scientific and Industrial Research, New Delhi on grant nos. 22(229) SP/92 EMR-II and 22(254)/96 EMR-II.

We would like to thank Mr G A Rakesh for his help in fabricating the boundary layer plate and for making the pitot-tube measurements. Lt A Bose, Mr Sharad K Gupta and Mr K Koren also helped in the measurements. The cooperation of Mr T R Bhogal and the other staff of the Gas Dynamics Laboratory is also gratefully acknowledged. Finally, we offer grateful thanks to Professor R Narasimha, FRS, for his constant encouragement, support and guidance.

References

Coles D F 1962 Rand Report R-403-PR, Rand Corp., Santa Monica, CA

Gaster M, Kit E, Wygnanski I 1985 Large-scale structures in a forced turbulent mixing layer. *J. Fluid Mech.* 150: 23–39

Hinze J O 1975 *Turbulence* 2nd edn (New York: McGraw-Hill)

Hussain A K M F, Reynolds W C 1970 The mechanics of an organised wave in turbulent shear flow. *J. Fluid Mech.* 41: 241–258

Hussain A K M F, Reynolds W C 1972 The mechanics of an organised wave in turbulent shear flow. Part 2. Experimental results. *J. Fluid Mech.* 54: 241–261

Hussain A K M F, Reynolds W C 1975 Measurements in fully developed turbulent channel flow. *J. Fluids Engg* 97: 568–578

Jang P S, Benney D J, Gran R L 1986 On the origin of streamwise vortices in a turbulent boundary layer. *J. Fluid Mech.* 169: 109–123

Klebanoff P S 1954 Characteristics of turbulence in a boundary-layer with zero pressure gradient. NACA technical report No. 3178

Landau L D 1944 On the problem of turbulence. *C. R. Acad. Sci. (URSS)* 44: 311–314

Liu J T C 1988 Contributions to the understanding of large-scale coherent structures in developing free turbulent shear flows. *Adv. Appl. Mech.* 26: 183–309

Malkus W V R 1956 Outline of a theory of turbulent shear flow. *J. Fluid Mech.* 1: 521–539

Patel V C 1965 Calibration of the Preston tube and limitations on its use in pressure gradients. *J. Fluid Mech.* 23: 185–208

Pope S B 1975 A more general effective viscosity hypothesis. *J. Fluid Mech.* 72: part 2: 331–340

Purtell L P, Klebanoff P S, Buckley F T 1981 Turbulent boundary-layer at low Reynolds number. *Phys. Fluids* 24: 802–811

Reynolds W C, Hussain A K M F 1972 The mechanics of an organized disturbance in turbulent shear flow. Part 3. Theoretical models and comparisons with experiments. *J. Fluid Mech.* 54: 263–288

Reynolds W C, Tiederman W G 1967 Stability of turbulent channel flow, with application to Malkus's theory. *J. Fluid Mech.* 27: 253–272

Roshko A 1992 Instability and turbulence in shear flows. *Theoretical and applied mechanics* (eds) S R Bodner, J Singer, A Solan, Z Hussain (Amsterdam: Elsevier)

Sen P K 1993 Some recent developments in the theory of boundary layer instability, *Sādhanā* 18: 387–403

Sen P K, and Vashist T.K 1989 On the nonlinear stability of boundary-layer flow over a flat plate. *Proc. R. Soc. London* A424: 81–92

Sen P K , Venkateswarlu D, Maji S 1985 On the stability of pipe-Poiseuille flow to finite-amplitude axisymmetric and non-axisymmetric disturbances. *J. Fluid Mech.* 158: 289–316

Sen P K, Rakesh G A, Veeravalli S V 1993 Analysis of wall turbulence using hydrodynamic stability theory. *Proc. 20th National Conf. on Fluid Mechanics and Fluid Power*, Palghat, B3-1 to B3-6

Sen P K, Veeravalli S V 1994 Hydrodynamic stability theory applied to wall turbulence. *Proc. 21st National Conf. on Fluid Mechanics and Fluid Power, Osmania University*, Hyderabad, pp TF-1 to TF-6

Sen P K, Veeravalli S V 1996a Evolution of an organised disturbance in a turbulent boundary layer. *Proc. 23rd National Conf. on Fluid Mechanics and Fluid Power, MACT*, Bhopal, pp 1–12

Sen P K, Veeravalli S V 1996b On a method of generating an analytically continuous turbulent mean velocity profile in turbulent boundary layer flow. *Proc. 23rd National Conf. on Fluid Mechanics and Fluid Power, MACT*, Bhopal, 13–18

Spalart P R 1988 Direct Numerical simulation of a turbulent boundary layer up to $R_\theta = 1410$. *J. Fluid Mech.* 187: 61–98

Tennekes H, Lumley J L 1972 *A first course in turbulence* (Cambridge, MA: MIT Press)

Townsend A A 1976 *The structure of turbulent shear flow* 2nd edn (Cambridge: University Press)

A fully automatic 3-D analysis tool for expansion chamber mufflers

R SRINIVASAN and M L MUNJAL

Department of Mechanical Engineering, Indian Institute of Science, Bangalore 560 012, India
e-mail: munjal@mecheng.iisc.ernet.in

MS received 15 April 1997; revised 7 August 1997

Abstract. The matrix condensation technique in conjunction with the substructuring principle has been previously used to model complex commercial automotive mufflers. Though complete 3-D finite element modelling is partially eliminated, it still requires the nodal-coordinate data and connectivity data of the elements forming the segment of each substructure. It also demands the connectivity of the degrees of freedom left after matrix condensation. Thus, the data preparation phase is tedious and cannot be claimed to be simple. Moreover, the existing FEM code requires the user to have a knowledge of acoustic theory and finite element modelling. This paper describes a way to completely eliminate the data preparation phase. Given the geometric dimensions of the muffler as input, the muffler performance is obtained as the output. The results obtained are compared with analytical and experimental results for simple as well as extended-tube expansion chambers, with or without an offset.

Keywords. Mufflers; duct acoustics; automatic mesh generation; finite element methods; noise control.

1. Introduction

Exhaust noise is a predominant source of noise in an automobile. Various types of muffler components are used to bring down the exhaust noise levels. Commercial exhaust mufflers have large cross-sections and complicated shapes. Therefore, 1-D analysis is not adequate for proper understanding of muffler performance. Thus, there is a need for a tool to perform a complete 3-D analysis of mufflers. The automotive industry is vibrant and fast-changing. Therefore, analysis tools used must be highly user-interactive and should be engineered to offer readymade solutions. The need for expertise on the part of the user should be eliminated so that industries need not rely upon experts to do muffler analysis. Thus, there is a need for a tool which will give muffler performance, given the geometric dimensions of the muffler.

While generally a one-dimensional (or, plane-wave) analysis (Munjal 1987a) is adequate, higher order modes do play an important role at higher frequencies in and around expansion chambers. Several methods have been proposed by the previous researchers for 3-D analysis of mufflers (El-Sharkawy & Nayfeh 1978; Ih & Lee 1985; Munjal 1987b; Abom 1990; Sahasrabuddhe *et al* 1991). Analytical solutions proposed by researchers (El-Sharkawy & Nayfeh 1978; Ih & Lee 1985; Abom 1990) could be one possible solution. However, these methods cannot yield a completely automatic code as they require eigenfunctions which are not readily known for complicated shapes. Therefore, one has to resort to methods not involving eigenfunction expansions. Various numerical techniques proposed in the literature were the collocation method, finite difference method and the popular finite element method. The collocation method (Munjal 1987) is relatively simple in-so-far as the algebra and computational effort are concerned, but is deficient in dealing with fractional area ratios and complicated shapes. Therefore, the collocation method fails to be a solution for the problem on hand. Finite difference method is not suitable for very complicated and irregular geometries (Baumeister 1980). However, the finite element method is completely general as far as the complex geometry of the muffler and different properties of the medium are concerned, and hence can handle rigid/non-rigid walled chambers, and simple or extended tube chambers with inline/offset inlet/outlet pipes (El-Sharkawy & Nayfeh 1978; Zienkiewicz 1979; Baumeister 1980; Ih & Lee 1985; Munjal 1987; Abom 1990; Sahasrabuddhe *et al* 1991).

Finite element method was extended to acousto-structural analysis by Gladwell (1965). Later, Young & Crocker (1975) and Craggs (1976) used FEM for prediction of muffler performance. Young & Crocker (1976) used rectangular Hermitian elements with C_1 continuity while Craggs used the axisymmetric ring element obtained from the hexahedral element. Their FEM predictions compared well with experiments. Later, Christiansen & Krenk (1988) used a recursive finite element technique for modelling pipes and ducts. They used an axisymmetric quadrilateral element and hence their analysis is restricted to 2-D configurations. Recently, Craggs (1989) analysed branched duct networks using FEM. A truly 3-D analysis of expansion chamber mufflers incorporating the recursive FEM technique of Christiansen & Krenk (1988) was carried out by Sahasrabudhe *et al*, (1991) who used the sub-structuring principle in conjunction with the matrix condensation technique for modelling simple/extended tube expansion chamber mufflers. Though Sahasrabudhe *et al* (1991) could analyse simple/extended tube expansion chambers, the FEM code was not user-interactive. The code requires the nodal coordinate data of the segments of the substructures, elemental connectivity data, and degrees of freedom left after condensation in addition to the dimensions of the muffler. Thus the data preparation phase is tedious and cumbersome. Moreover, the user should have expertise in acoustic theory, finite element modelling etc. Thus the need for a truly 3-D analysis tool is not fulfilled.

In this paper, the sub-structuring and the matrix condensation technique have been used to model muffler geometry. Given the dimensions of the muffler as the input, the 3-D finite element model is generated and muffler performance is obtained as the output. The results obtained using the automatic code are compared with the results available in the

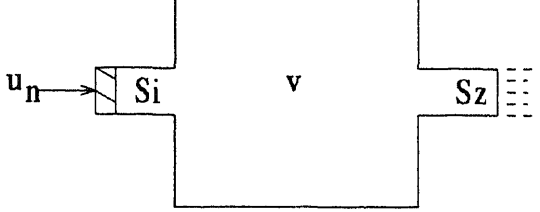


Figure 1. Acoustic system.

2. Finite element formulation

The acoustic field in a muffler configuration filled with stationary ideal fluid (figure 1) is given by

$$[\nabla^2 + k^2]p = 0, \tag{1}$$

with the following four types of boundary conditions:

(1) $$\partial p/\partial n = -j\rho_o\omega(p/Z), \tag{2}$$

when the normal acoustic impedance Z is known;

(2) $$\partial p/\partial n = -j\rho_o\omega u_n, \tag{3}$$

when the normal particle velocity is known as in the case of the constant velocity piston source;

(3) $$\partial p/\partial n = 0, \tag{4}$$

at the rigid walls of the muffler; and

(4) $$p = p_n, \tag{5}$$

when the constant pressure source is assumed.

Assuming a trial solution for the field variable p, at any point within an element as

$$p = \{N\}_e^T\{p_n\}_e, \tag{6}$$

where N_e is the basis function or the shape function or the interpolation function. Substituting (6) into (1), we get the residual as

$$Re = [\nabla^2 N_e^T + k^2 N_e^T]p_{ne}. \tag{7}$$

This error is orthogonalized with respect to the weighting functions to get

$$\int_{V_e} N_e Re\, dv = 0, \tag{8}$$

which can be written as

$$\int_{V_e} N_e \nabla^2 N_e^T p_{ne} dv + \int_{V_e} k^2 N_e N_e^T p_{ne} dv = 0. \tag{9}$$

Then, following Munjal (1987a), we get

$$\left[[M]_e - k^2[S]_e + \frac{j\rho_o\omega}{z}[D]_e\right] p_{n_e} = -j\rho_o\omega F_e, \tag{10}$$

$$[M]_e = \int_{V_e} \nabla N_e \nabla N_e^T \, dV, \quad \text{(mass matrix)}, \tag{11}$$

$$[S]_e = \int_{V_e} N_e N_e^T \, dV, \quad \text{(stiffness matrix)}, \tag{12}$$

$$[D]_e = \int_{S_e} N_e N_e^T \, dS, \quad \text{(damping matrix)}, \tag{13}$$

and,

$$\{F\}_e = \int_{S_e} u_n N_e dS, \quad \text{(forcing vector)}. \tag{14}$$

Here, $[S]$ denotes the stiffness matrix and the surfaces S_{e_z}, S_{e_i} and S_{e_r} denote the portion of the boundary in which normal acoustic impedance, the input source velocity and the rigid wall boundaries are known, respectively. Surface S_e includes S_{e_z}, S_{e_i} and S_{e_r}. Equation (10) can be written after global assembly as,

$$[K]\{p\} = \{F\} \tag{15}$$

where $[K]$ is called the admittance matrix given by

$$[K] = \sum_{i=1}^{l} [K]_e = \sum_{i=1}^{l} \left[[M]_e - k^2[S]_e + \frac{j\rho_o\omega}{z}[D]_e\right], \tag{16}$$

and

$$\{F\} = \sum_{i=1}^{l} \{F\}_e = -j\rho_o\omega \sum_{i=1}^{l} \{F\}_e, \tag{17}$$

where l denotes the number of elements in the system. The contribution to forcing vector $\{F\}$ comes only from the surface S_i at the inlet of the muffler assuming constant velocity source and contribution to the damping matrix $\{D\}$ comes from the only surface Sz at the outlet of the muffler (see figure 1).

3. Muffler performance using FEM

Higher order modes or 3-D effects are generally limited to expansion chambers and a couple of diameters length of the inlet tube and outlet tube respectively. Therefore, a muffler composed of a volume chamber with inlet and outlet tubes (figure 2) can be simulated by a linear acoustic four-pole system (Munjal 1987a). Making use of this analogy, the acoustic state variables viz., acoustic pressure and volume velocity on the inlet side can be related to those on the outlet side through the four-pole parameters also called the transfer matrix. The transfer matrix relation can be written as

$$\begin{bmatrix} p_n \\ v_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} p_1 \\ v_1 \end{bmatrix}$$

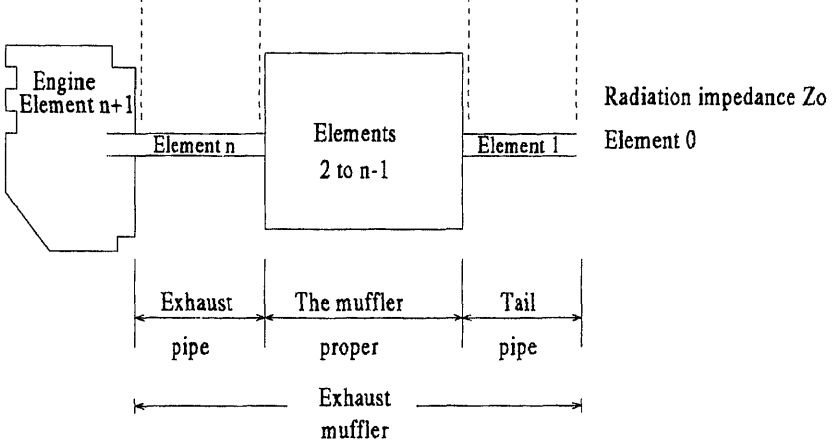


Figure 2. Engine exhaust system.

where p and v are, respectively, acoustic pressure and volume velocity, subscripts n and 1 stand for exhaust pipe and tail pipe (figure 2) respectively, and four-pole parameters are given by

$$A_{11} = \left[\frac{p_n}{p_1}\right]_{v_1=0}, \tag{18}$$

$$A_{12} = \left[\frac{p_n}{v_1}\right]_{p_1=0}, \tag{19}$$

$$A_{21} = \left[\frac{v_n}{p_1}\right]_{v_1=0}, \tag{20}$$

$$A_{22} = \left[\frac{v_n}{v_1}\right]_{p_1=0}. \tag{21}$$

The assembled set of dynamic equations, (10), obtained through finite element formulation, are solved twice using two sets of boundary conditions to calculate four-pole parameters.

(1) By imposing $v = 0$ $(i.e. \partial p/\partial x = 0)$ at the outlet boundary, nodal points A_{11} and A_{21} are computed using (18) and (20).

(2) By imposing $p = 0$ at the outlet boundary, nodal points A_{12} and A_{22} are computed using (19) and (21).

Then the transmission loss (TL) value can be evaluated as,

$$\mathrm{TL} = 20 * \log\left[\frac{1}{2} * |A_{11} + S_1\frac{A_{12}}{\rho_o c_o} + \frac{A_{21}\rho_o c_o}{S_n} + A_{22}|\right], \tag{22}$$

where S_n and S_1 are the areas of the exhaust pipe and tail pipe respectively, and v is the acoustic volume velocity,

ρ_o = density of the medium,

and c_o = sound speed.

4. Selection of the finite element for muffler analysis

In general, there is a wide-choice of finite elements available in the literature. However, the 20-noded isoparametric Brick element (Zienkiewicz 1979) has been used extensively throughout this work. The reasons for selecting Isoparametric 20-noded brick elements are as follows.

(1) Since automotive mufflers involve curved geometry, 20-noded isoparametric elements with quadratic shape functions are found suitable. A large number of linear elements may be required to discretize the same geometry. Therefore, use of a higher-order element reduces memory requirement and solution time.

(2) 20-noded isoparametric elements have been successfully used by Sahasrabudhe *et al* (1991) for finite element modelling of mufflers (Munjal 1987a) and the results obtained are satisfactory.

5. Maximum length of a segment and element

Investigation of the criteria for the maximum length of a typical segment is presented here. Ross (1981), based on his numerical experiments, reported that, in order to ensure convergence of finite element results up to a frequency of f-max, the maximum typical dimension of the finite element should be less than 0.2 times the minimum wavelength. He used 8-noded isoparametric elements in his mesh configuration. Later, Sahasrabudhe *et al* (1991) reported that, with the use of a quadratic 3-D element (20-noded isoparametric element), the maximum typical dimension of the finite element should be less than 0.52 times the minimum wavelength. To verify this criterion, an analytical solution was derived for a simple square pipe and it was compared with the results using FEM.

Details of the comparison are as follows:

A square pipe with its boundary conditions is shown in figure 3. The dimensions of the pipe are such that only plane waves can propagate. Therefore, the governing equation of the problem is

$$(D^2 + k^2)p(z) = 0, \tag{23}$$

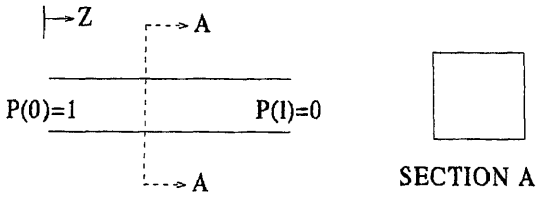


Figure 3. Square pipe model.

and the boundary conditions are:

- @ $z = 0$, $p(0) = 1$,
- @ $z = 1$, $p(1) = 0$.

The solution for (23) is

$$p(z) = C1 \cdot \exp(-jkz) + C2 \cdot \exp(+jkz), \tag{24}$$

where $C1$ and $C2$ are constants and k is the wave number. The constants $C1$ and $C2$ can be found by applying the two boundary conditions. Thus, the solution is given by

$$p(z) = \sin k(l + z) \sin(kl), \tag{25}$$

where

$$l = (n \cdot c_o)/(4 \cdot f), \quad \text{for} \quad n = 1, 3, 5, 7 \cdots. \tag{26}$$

Equation (25) is taken as a reference for validation of the FEM results.

To verify the criteria for maximum element length, the following test runs were conducted.

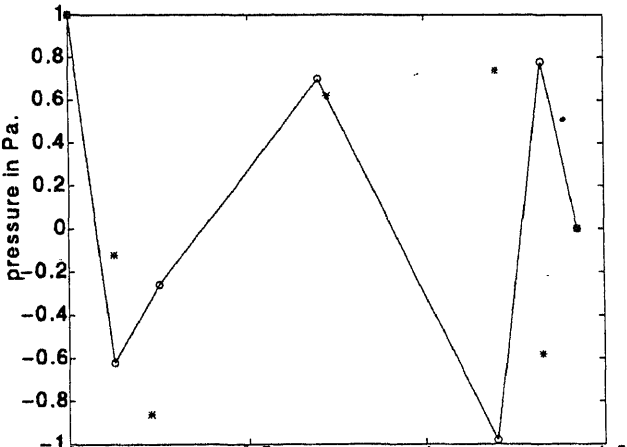
Data for the test run:

(1) Frequency 1000 Hz,

(2) length $l = (n \cdot c_o)/(4 \cdot f), n = 1, 3, 5, \ldots,$

(3) side of the square $h < c_o/(2 \cdot f)$

The following procedure was adopted for conducting the test, and the steps involved are as follows:



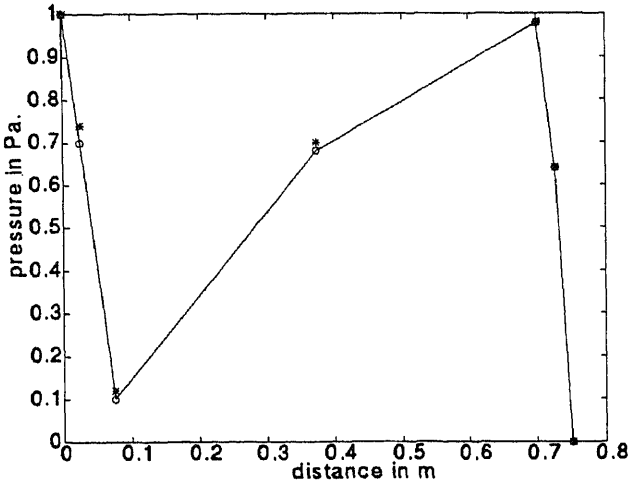


Figure 5. Comparison of FEM predictions (∗) for a square pipe model of length 0.765 m, side 0.05 m and a segment length 0.225 times the wavelength with analytical solution (○) given by (25).

- Select a suitable value of the pipe length using (26),
- vary the segment length each time (greater than 0.52 times the minimum wavelength to start with),
- the value of h (side of the square cross-section) is taken equal to the segment length so that each element has an aspect ratio of unity,
- compare the results with the analytical solution.

Analysis was conducted with a segment length equal to 0.7, 0.54, 0.375, 0.225 times the minimum wavelength. The results for two cases are shown in figures 4 and 5. The solution obtained using 0.225 times the minimum wavelength (figure 5) tallied quite well with the analytical solution. From the above results it may be observed that the element length must be certainly less than 0.375 times the minimum wavelength.

A limitation of the foregoing analysis is that it is conducted in the plane wave zone where the effect of higher order modes is not present. Hence, to check the validity of the above tests further, a complete 3-D analysis of a simple expansion chamber muffler was performed and the computed values of TL were compared with those obtained from an analytical solution. The tests are as below.

A simple expansion chamber with a diameter ratio of 3 and length to diameter ratio of 1.67 was analysed and the transmission loss thus obtained was compared with the results available in the literature (El-Sharkawy & Nayfeh 1978). The segment length was taken as 0.125 m. From figure 6, it may be observed that the results agree well up to a frequency of 865 Hz. This observation leads to the conclusion that the segment length should be less than or equal to 0.3 times the minimum wavelength of interest. Figure 7 shows the comparison between TL of a simple expansion chamber with the same dimension as in the previous case but differing in that the segment length taken is 0.062 m. The results agree quite well up to a non-dimensional frequency parameter kl_2 of 8 which corresponds to

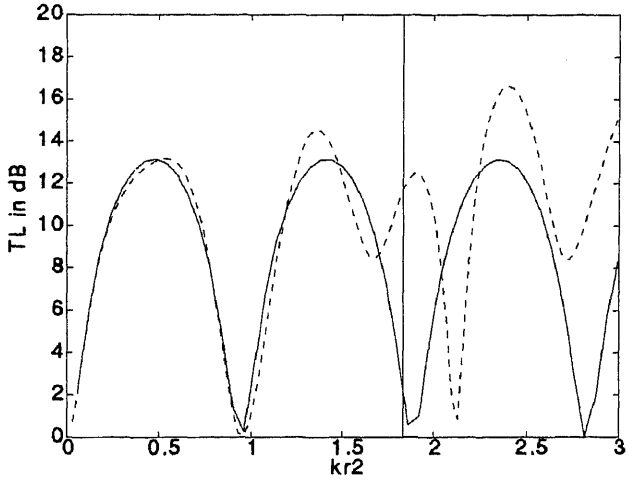


Figure 6. Comparison of TL of a coaxial expansion chamber with $l_1 = l_3 = 0.375$ m, $d_1 = d_3 = 0.05$ m, $l_2 = 0.2505$ m, $d_2 = 0.15$ m, and with segment length 0.125 m, (- - -) FEM, (—) analytical (El-Sharkawy & Nayfeh 1978).

a frequency of 1730 Hz. This observation confirms the previous conclusion. From these numerical experiments the following conclusions were derived.

- A 20-noded isoparametric element can be used to model pipes with curved geometry.
- Maximum typical dimension has to be less than 0.3 times the minimum wavelength when using the 20-noded isoparametric element and less than 0.2 times the minimum wavelength when using linear elements as given by Ross (1981).

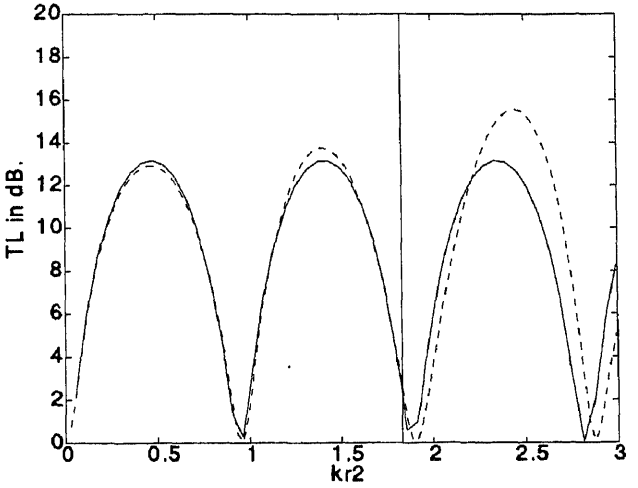


Figure 7. Comparison of TL of a coaxial expansion chamber with $l_1 = l_3 = 0.187$ m, $d_1 = d_3 = 0.05$ m, $l_2 = 0.2505$ m, $d_2 = 0.15$ m, and with segment length 0.062 m, (- - -) FEM, (—) analytical (El-Sharkawy & Nayfeh 1978).

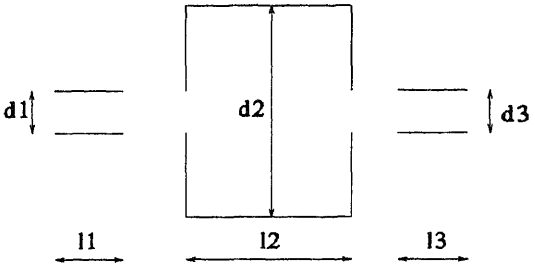


Figure 8. Simple expansion chamber.

6. Modelling of simple expansion chambers

A typical simple expansion chamber muffler is shown in figure 8.

Basically it consists of three substructures (Munjal 1987a):

- Inlet pipe (substructure 1);
- expansion chamber (substructure 2); and
- the outlet pipe (substructure 3).

The task of a mesh generator is to discretize the volume of the muffler. The FEM code takes the length and diameter of the inlet pipe, the expansion chamber and the outlet pipe respectively as the inputs. The basis for the finite element modelling is the matrix condensation technique (Munjal 1987a). The construction of a 3-D finite element model proceeds as follows.

Nodes are generated in a 2-D region as shown in figure 9 which uses only the inlet pipe diameter and the chamber diameter. The mesh shown is a cross-section of a typical segment of substructure 2. This 2-D mesh can be extruded to form a typical segment of substructure 2 (Sahasrabudhe *et al* 1991). The 8-nodes on the periphery of the inner circle are extruded to form pipe segments for the inlet/outlet pipes. Number of segments of substructure 2 is given by m where it becomes $m + 1$ for inlet/outlet pipes. The segment at the inlet section of substructure 1 and outlet section of substructure 3 are not condensed, as the gradient of pressure is computed using these segments. This is useful in the computation of four-pole parameters as discussed.

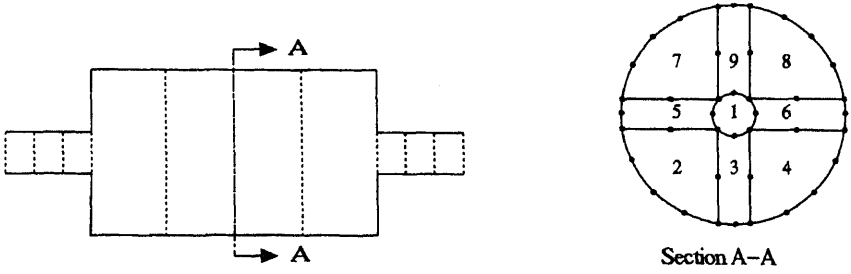


Figure 9. Simple expansion chamber discretization.

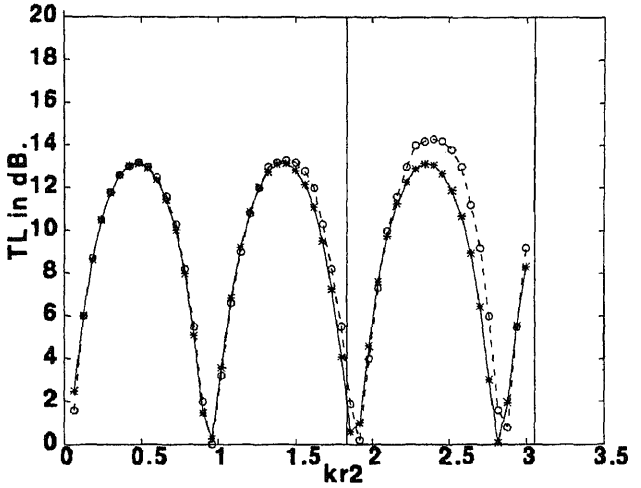


Figure 10. Comparison of FEM results for a simple expansion chamber with $d_1 = 0.05$ m, $d_2/d_1 = 3, l_2/d_2 = 1.67$ with the analytical solution of El-Sharkawy & Nayfeh (1978). (*) analytical, (o) FEM.

Here, m is any positive integer such that the ratio $1/2^m$ is less than 0.3 times the inimum wavelength. One has to note that only nine elements are generated irrespective f the muffler dimension. This mesh is found to be sufficient to analyse mufflers of practical terest. Since the same number of elements are formed, the connectivity data for a typical gment remain the same and are read from a data file. In the same manner, the information garding the degrees of freedom left after matrix condensation is read from the data file. hus, the geometric dimensions alone are sufficient to obtain a 3-D finite element model f a simple expansion chamber muffler. Limitations of the code are as follows.

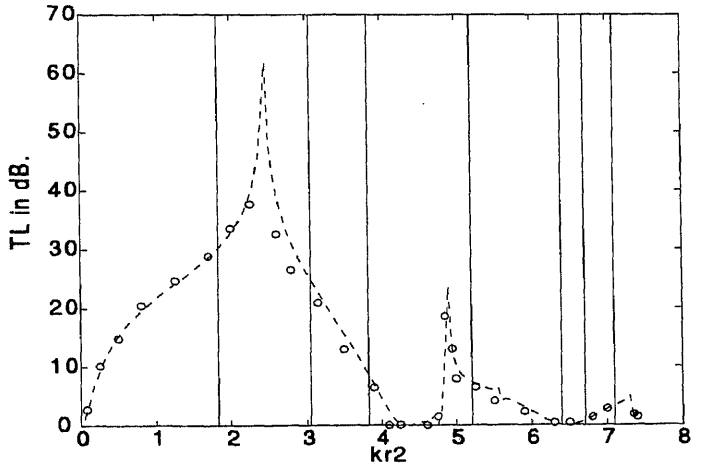


Figure 11. Comparison of the TL for a simple expansion chamber with $d_1 = 0.0205$ m,

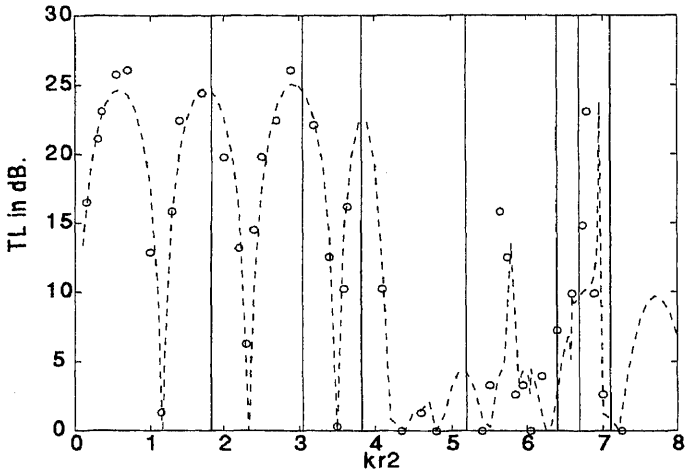


Figure 12. Comparison of the FEM predictions for coaxial expansion chamber with $d_1 = 0.0205$ m, $d_2/d_1 = 6$, $l_2/d_2 = 1.35$ with the experimental results of Ih & Lee (1985) (- - -) FEM, (o) experiments.

- The circumference of the inlet/outlet pipe should be less than 1.2 times the minimum wavelength of interest.
- The difference between the radius of inlet pipe and expansion chamber should not exceed 0.3 times the minimum wavelength.

These limitations may easily be overcome by using an adaptive mesh generator which will create more elements for a bigger dimension so that the maximum typical dimension of any of the elements does not exceed the prescribed limit. Thus the process of finite element modelling is fully automated without user intervention.

To validate the modelling procedure discussed in the previous section, the results obtained using the completely automatic FEM code were compared with the experimental and analytical results available in the literature. They are as shown in the figures 10–12. The finite element results showed good agreement with the analytical solution (El-Sharkawy & Nayfeh 1978) throughout the frequency range of interest as can be seen in figure 10. From figures 11 and 12, it is seen that the FEM predictions are in good agreement with the experimental results as well.

7. Modelling of extended tube expansion chambers

A typical extended tube expansion chamber with five substructures is shown in figure 13. One may note that modelling of extended tube expansion chambers is very similar to that of simple expansion chambers. The difference lies in the additional task of modelling the two additional substructures (annular cavity introduced by the extended pipes). Figure 13 shows the discretized cross-section of the annular cavity between the inlet pipe and the chamber. The nodes are formed in much the same way as the simple expansion chamber but the difference lies in the connectivity. In the annular cavity model, the element representing the cross-section of the inlet pipe is absent. It is to be noted here that there is no interaction

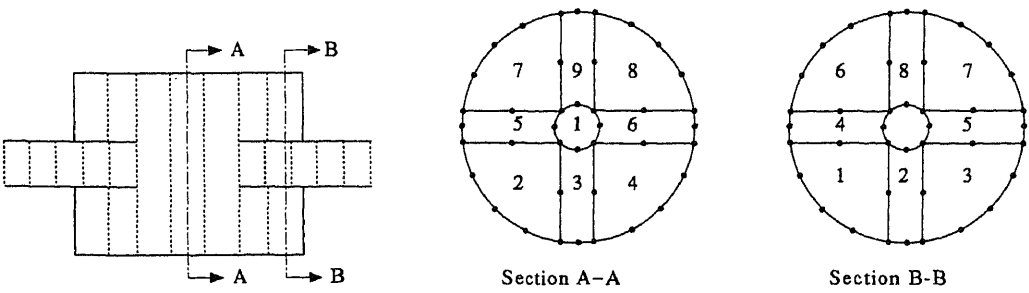
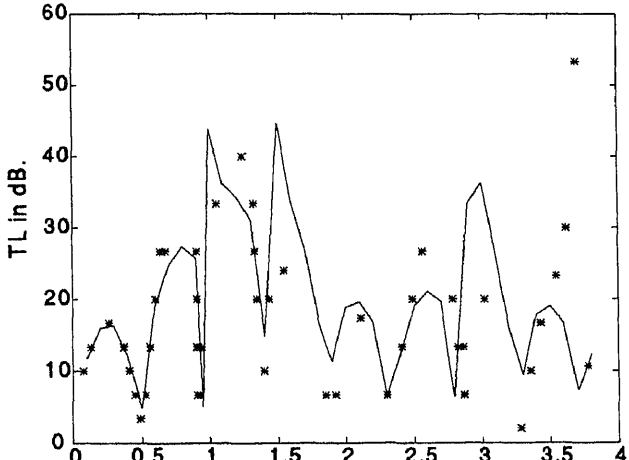


Figure 13. Discretized cross-sections of the annular cavity and the expansion chamber.

between the substructures 1 and 2, or 4 and 5 as the walls of the pipe are assumed to be rigid. The interaction is only through substructure 3. The input to the FEM code will be the length and diameter of the five substructures. The FEM program can handle the following cases:

- extended inlet and extended outlet,
- simple inlet and extended outlet,
- extended inlet and simple outlet.

To validate the modelling procedure illustrated above, an extended tube expansion chamber was analysed and the results were compared with the experimental results of Ih & Lee (1985). The comparison is shown in figure 14.



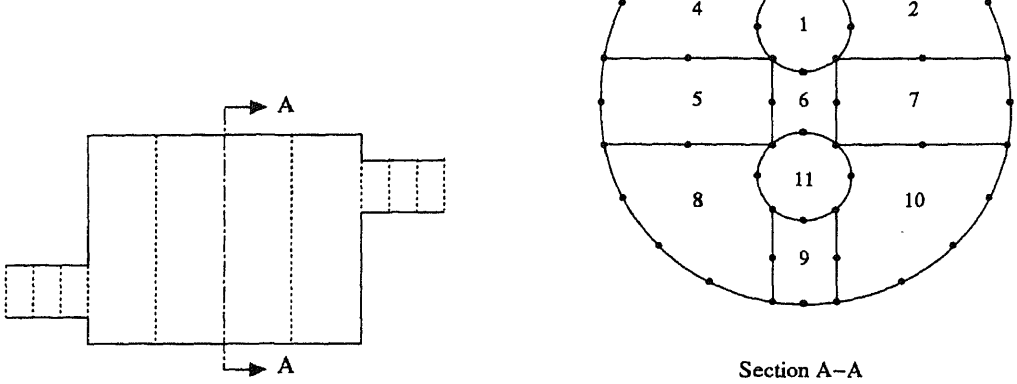


Figure 15. Discretized cross-section of an offset inlet/outlet model.

8. Modelling of offset inlet/outlet chambers

In the case of offset pipe chambers, the automatic FEM code requires as the input the centre and radii of the pipes forming the substructures. The mesh generation routines are capable of handling any arbitrary position of the inlet and outlet pipes provided the two pipes do not intersect. Figure 15 shows the mesh generated for an angular position of 180° between the inlet and outlet pipes.

Two-dimensional nodal data are converted into a segment model as shown in figure 15, which after condensation is representative of the whole volume of the muffler. Thus

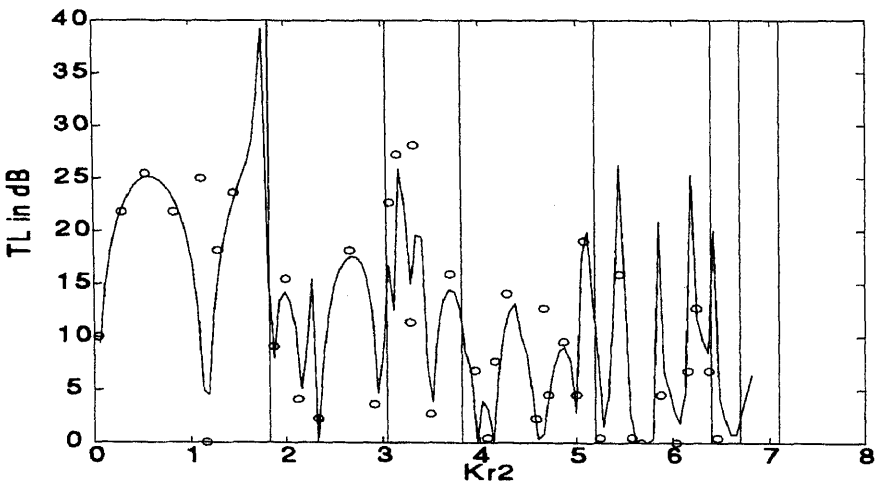


Figure 16. Comparison of the FEM predictions (—) for an offset expansion chamber with the experimental results of Ih & Lee (1985). $\alpha = 6$, $\beta = 1.35$; $r_1 = r_3 = 10.25$ mm; $\delta_1 = \delta_2 = 0.37r_2$; $\theta = 180°$.

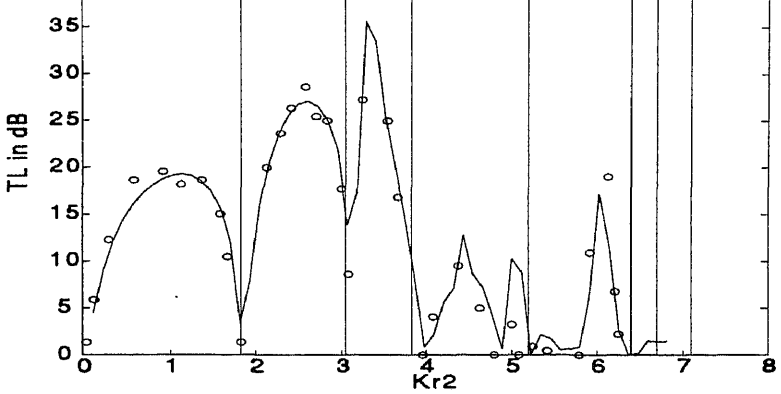


Figure 17. Comparison of FEM predictions (—) for an offset expansion chamber with the experimental results of Ih & Lee (1985). $\alpha = 6$, $\beta = 0.33$; $r_1 = r_3 = 10.25\,\text{mm}$; $\delta_1 = \delta_2 = 0.37r_2$; $\theta = 180°$.

matrix condensation eliminates the task of constructing a complete 3-D model, which is quite cumbersome and consumes a lot of memory and CPU time. The nodal connectivity data for the segments of different substructures are read from a data file. The number of segments in each of the substructures is calculated in such a way that the ratio $l/2^m$ (segment length) is less than 0.3 times the minimum wavelength. Here m denotes the number of successive matrix condensations (Christiansen & Krenk 1988). Thus the process of modelling becomes fully automatic.

Figures 16 and 17 corroborate the finite element results obtained in the form of TL of acoustically long as well as short offset inlet/outlet expansion chamber mufflers with the experimental results of Ih & Lee (1985).

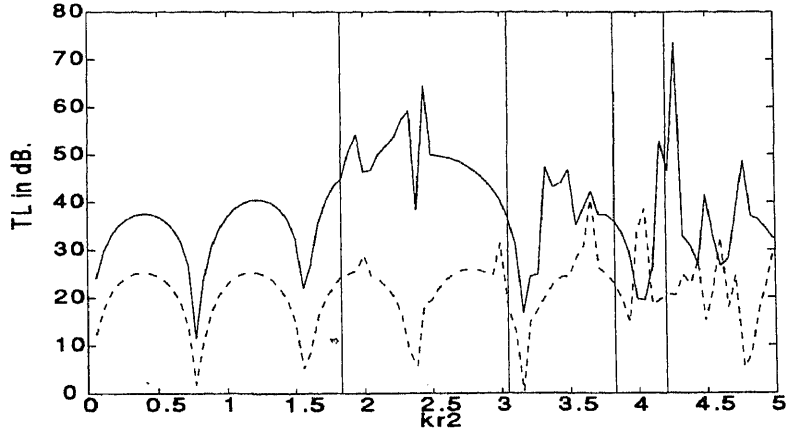


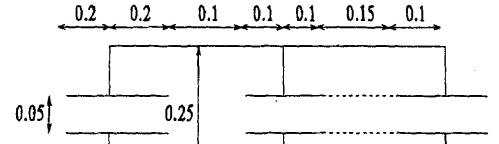
Figure 18. Comparison of the TL of the offset simple (- - -) and offset extended tube (—) expansion chambers with l_{ei} and $l_{eo} = 0.33d_2$ for the extended tube. $\alpha = 6$, $\beta = 2$; $\delta_1 = 0$, $\delta_2 = 0.63r_2$, $\theta = 0°$.

9. Modelling of offset extended inlet/outlet pipes

The modelling of mufflers with inlet/outlet pipes offset and extended into the expansion chamber is done using the concepts discussed in the previous sections. Two more sub-structures are introduced by extending the inlet/outlet pipes. These cavities are modelled in the same way as the offset expansion chamber, the difference being the absence of the element corresponding to the pipe segment. While the element eleven corresponding to the inlet pipe is absent in the segment model of substructure two, element one is missing in the segment model of the substructure four (see figure 15). A typical muffler with axial inlet and radially offset outlet pipe was analysed with and without the extension of the inlet/outlet pipes and the results compared. The comparison is shown in figure 18. The resonances in the end chambers are found to improve the performance of the extended tube chambers as compared to their simple chamber counterpart.

10. Integration of 3-D analysis with 1-D analysis

Often, higher order modes that are generated in expansion chambers decay quickly in the adjoining inlet/outlet pipes, where only plane waves can propagate. Thus, the completely automatic 3-D analysis subroutines developed for an expansion chamber with small (about two diameters long) portions of the inlet pipe and outlet pipe can be used to generate a 2×2 transfer matrix as described above. This can then be integrated with Munjalś Transfer Matrix-based Muffler Program (TMMP) which, in fact, was the primary purpose of the investigation. Whenever expansion chambers appear as a part of the acoustic system, the 3-D analysis routines are called upon to get the four-pole parameters connecting the inlet and outlet sections of the chamber. Thereafter, respective routines of the program are used to perform plane-wave analysis. The overall performance of the acoustic system can be obtained by simply multiplying the successive transfer matrices (Munjal 1987a). As an illustration, consider an acoustic system which consists of a simple expansion chamber and a concentric tube resonator in series as shown in figure 19. The overall performance of the system is calculated and plotted in figure 20. It may be noted that 1-D analysis results show good agreement up to the cut-on of the first azimuthal mode, but show discrepancies in the high frequency range. This observation clearly illustrates the inadequacy of plane-wave analysis above the cut-on frequencies of higher order modes. But the CPU time for each frequency step in the case of a 3-D FEM program is at least 100 times that of a plane-wave analysis program. Therefore, one has to resort to 3-D analysis only when the higher order mode effects predominate within the frequency range of interest. Thus it is possible to analyse any combination of simple/extended tube chambers in series using the TMMP



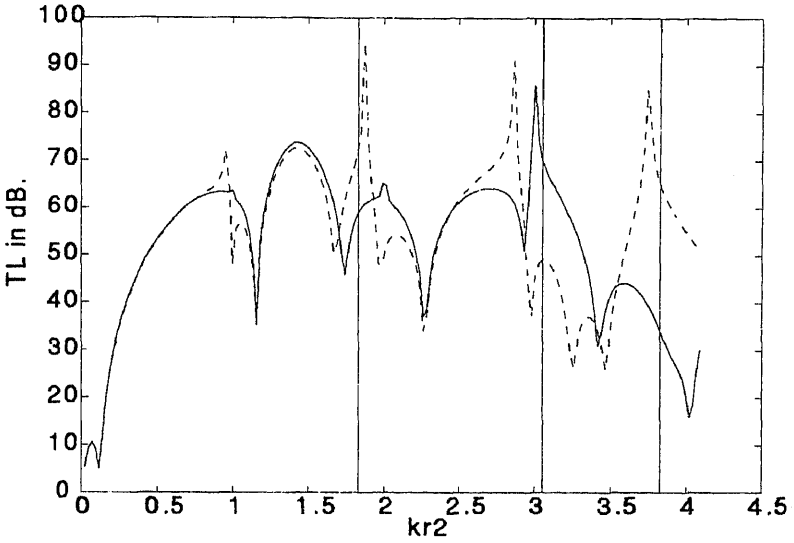


Figure 18. Comparison of FEM predictions (- - -) with TMMP (—) for the acoustic system shown in figure 19.

with 3-D analysis routines being restricted to chambers whenever the user feels that 3-D analysis is inevitable.

11. Concluding remarks

- Maximum typical dimension of a finite element should be less than 0.3 times the minimum wavelength of interest while using quadratic shape functions and 0.2 times the minimum wavelength for linear shape functions, as suggested by Ross (1981).
- A completely automatic FEM program which takes dimensions of the muffler as the input and four pole parameters as the output has been developed for analysing simple/extended tube expansion chambers with coaxial or offset location of the inlet/outlet pipes and integrated with Munjal's TMMP software. This work thus fulfills the need for a user-friendly tool for 3-D analysis. It is to be noted that the need for an expert with knowledge of acoustic theory and finite element modelling is totally eliminated.

References

Abom M 1990 Derivation of four-pole parameters including higher order mode effects for expansion chamber mufflers with extended inlet and outlet. *J. Sound. Vib.* 137: 403–418

Baumeister K J 1980 Numerical techniques in linear duct acoustics. NAA-TM-81553, E-513

Christiansen P S, Krenk S 1988 A recursive finite element technique for acoustic fields in pipes with absorption. *J. Sound. Vib.* 122: 107–118

Craggs A 1989 The application of the transfer matrix and matrix condensation methods with finite elements to duct acoustics. *J. Sound. Vib.* 132: 393–402

El-Sharkawy I, Nayfeh A H 1978 Effect of an expansion chamber on the propogation of sound in circular ducts. *J. Acoust. Soc. Am.* 63: 667–674

Gladwell G M L 1965 *Proceedings of Fifth International Congress on Acoustics*, Liege, L-33.

Gopinathan V 1989 *3-D analysis of simple axisymmetric expansion chamber mufflers for stationary medium by the exact method.* M E dissertation, Indian Inst. Sci., Bangalore

Ih J G, Lee B H 1985 Analysis of higher-order mode effects in the circular expansion chamber with mean flow. *J. Acoust. Soc. Am.* 77: 1377–1388

Munjal M L 1987a *Acoustics of ducts and mufflers* (New York: John Wiley and Sons)

Munjal M L 1987b A simple numerical method for three-dimensional analysis of simple expansion chamber mufflers of rectangular as well as circular cross-section with stationary medium. *J. Sound Vib.* 116: 71–88

Ross D F 1981 A finite element analysis of parallel-coupled acoustic systems. *J. Sound. Vib.* 79: 133–143

Sahasrabudhe A D, Munjal M L, Anantharamu S 1991 Matrix condensation and transfer matrix techniques in the 3-D analysis of expansion chamber mufflers. *J. Sound Vib.* 147: 371–394

Young C I J, Crocker M J 1975 Prediction of transmission loss in mufflers by the finite element method. *J. Acoust. Soc. Am.* 57: 144–148

Zienkiewicz O C 1979 *The finite element method* 3rd edn (New Delhi: Tata McGraw-Hill)

Radiative heat transfer in participating media – A review

SUBHASH C MISHRA [1,*] and MANOHAR PRASAD [2]

[1]Department of Mechanical Engineering, Indian Institute of Technology, Guwahati 781 001, India
[2]Department of Mechanical Engineering, Indian Institute of Technology, Kanpur 208 016, India
e-mail: scm@iitg.ernet.in; manohar@iitk. ernet.in

MS received 26 August 1996; revised 13 April 1998

Abstract. This paper presents an overview of various exact analytic and approximate numerical methods for the solution of radiative heat transfer problems in participating media. Review of each method is followed by its strengths and limitations. Importance of radiative heat transfer analysis and difficulties in the solution of radiative transfer problems have been emphasized.

Keywords. Radiation; participating media; radiative transfer equation.

1. Introduction

The rapidly depleting fossil fuel reserves, the growing ecological imbalance due to combustion products, the requirements of the space age etc., have necessitated the designing of energy efficient systems, and this in turn requires accurate prediction of heat transfer rates in these systems. Of the three modes of heat transfer, conductive and convective heat transfer rates are generally linearly proportional to temperature differences, whereas radiative heat transfer rates are mostly proportional to differences in the fourth powers of the absolute temperatures. Thus, in high-temperature applications, determination of radiative heat transfer rate becomes an important and necessary consideration. Consideration of thermal radiation becomes paramount in the design of high temperature energy conversion devices (boilers, engines, rocket nozzles etc.), in industrial processes (heating furnaces, incinerators), in nuclear reactors (such as in the Sun, fusion reactors, nuclear bombs), during atmospheric reentry of space vehicles, in materials processing (glass, metals), in utilization of materials (metallic screens, ceramic foams), in analyses of industrial and building fires etc. Thus, this wide range of potential applications provides strong motivation for continued research in the field of radiative heat transfer.

*For correspondence

Especially in combustion applications, either for power production or propulsion, consideration of thermal radiation becomes extremely important because combustion phenomenon as well as heat transfer rates are greatly influenced by radiation. For example, in large-scale coal combustion chambers (Manickavasagam & Menguc 1993) and in fluidized bed combustion (Tien 1988), radiative heat transfer accounts for as much as 90% and 40% respectively of the total heat transfer rates. Further, accurate determination of radiative heat transfer is essential in the design of burners, furnaces, and internal combustion (IC) engines. In these systems, minimization of emission of the oxides of nitrogen, NO_x, is one of the primary concerns, because NO_x formation is highly sensitive to the changes in temperature (Kuo 1986).

Although in high temperature applications radiation remains the dominant mode of heat transfer, for evaluation of total heat transfer rate, conduction and convection also have to be considered. In the solution of any high temperature flow problem, which involves conduction, convection, and radiation (conjugate problem), apart from the usual difficulties encountered due to the large number of partial differential equations required for the solution of conductive and convective parts of the problem, a greater complexity in problem formulation as well as much higher computational cost become imminent due to the presence of an extra radiative heat flux term in the general equation for energy conservation.

In full simulation of the heat transfer process in a typical high-temperature energy conversion device, say an internal combustion engine, one has to consider three momentum equations, at least two turbulent equations, equations for species, and one general equation for energy conservation incorporating all the three modes of heat transfer. The general equation for energy conservation for moving compressible fluid is given as

$$\rho c_v \left(\frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T \right) = \underbrace{-\nabla \cdot (\underbrace{-k \nabla T}_{\text{conduction}} + \underbrace{\mathbf{q}_R}_{\text{radiation}})}_{\text{divergence of total heat flux}} - p \nabla \cdot \mathbf{v} + \mu \Phi + \dot{Q}'''. \tag{1}$$

In the general energy conservation equation (1), the term on the left hand side represents the contribution to energy transfer owing to convection, whereas the terms on the right hand side respectively represent contributions of conduction, radiation, pressure work, viscous dissipation, and internal generation.

Even though in (1), radiative heat flux is part of the total heat flux term, it itself is the most difficult part to evaluate, particularly in the presence of participating media. The radiative part has to be determined from the equation of conservation of radiative energy, which is an integro-differential equation, and is given as

$$\nabla \cdot \mathbf{q}_R = \kappa_\lambda [4\pi i_{\lambda b}(\tau_\lambda) - G(\tau_\lambda)], \tag{2}$$

where the blackbody intensity, $i_{\lambda b}$, in (2), is the Planck's intensity function which depends explicitly upon the local temperature, T, of the medium and κ_λ is the local linear spectral absorption coefficient of the medium. The incident radiation, G, in the above equation, is given by

$$G(\tau_\lambda) = \int_0^{2\pi} \int_0^{\pi} i(\tau_\lambda, \theta, \phi) \sin\theta \, d\theta \, d\phi, \tag{3}$$

where in the above equation, the intensity of radiation at any optical depth, τ_λ, in an absorbing, emitting, and scattering medium is given by

$$i(\tau_\lambda, \theta, \phi) = i(0, \theta, \phi) \exp(-\tau_\lambda) + \int_0^{\tau_\lambda} S(\tau_\lambda', \theta', \phi') \exp(-(\tau_\lambda - \tau_\lambda')) \, d\tau_\lambda'. \tag{4}$$

Equation (4) is the integral form of the radiative transfer equation (RTE). It describes the intensity of radiation at some optical depth τ_λ in the medium in terms of the intensities reaching that position from all other locations τ_λ' in the medium and from the medium boundary at $\tau_\lambda = 0$. It is obtained by integrating the differential form of the RTE (given below) from the point of origin of the ray, $\tau_\lambda' = 0$ at the wall, to a point $\tau_\lambda' = \tau_\lambda$ inside the enclosure.

$$\frac{di(\tau_\lambda, \theta, \phi)}{d\tau} + i(\tau_\lambda, \theta, \phi) = S(\tau_\lambda, \theta, \phi). \tag{5}$$

The source function in (5) is given by

$$S(\tau_\lambda, \theta, \phi) = (1 - \omega) \, i_{\lambda b}(\tau_\lambda) + \left(\frac{\omega}{4\pi}\right) \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} i(\tau_\lambda, \theta', \phi') \times p((\theta', \phi') \rightarrow (\theta, \phi)) \sin\theta' \, d\theta' \, d\phi', \tag{6}$$

where the single scattering albedo, $\omega_\lambda = \sigma_{\lambda s}/(\kappa_\lambda + \sigma_{\lambda s})$ and $\sigma_{\lambda s}$ is the scattering coefficient, and $p((\theta', \phi') \rightarrow (\theta, \phi))$ is called the anisotropic scattering phase function which represents the probability distribution of the energy scattered into the direction (θ, ϕ) from the rays traversing the elemental volume from all other directions, (θ', ϕ').

It should be noted that the RTE given in (4) is a third-order integral equation in intensity, $i(\tau_\lambda, \theta, \phi)$. For its evaluation, the integral over the source function must be carried out over the optical coordinate, τ_λ. But the source function itself is an unknown function of intensity (see (6)) which has to be integrated over two directions, θ and ϕ. Furthermore, usually the temperature T which is related to the blackbody intensity term, $i_{\lambda b}$, in the source function, (6), is not known and it must be found in conjunction with the overall conservation of energy, (1), thus further complicating the situation.

Therefore, due to the presence of radiation, the general equation for energy conservation, (1), turns out to be an integro-differential equation which poses enormous difficulties in its solution. Further, complexities due to radiation crop up due to the fact that radiative heat transfer deals with seven independent variables, viz., three space coordinates, the frequency of radiation, two coordinates describing the direction of photon travel, and time, whereas conduction and convection involve a maximum of four independent variables, viz., three space coordinates and time. Thus, the combination of the above complexities makes the complete solution of the problem, in the real sense, a formidable task even with the recent advent of supercomputers. Thus, there has always been the quest for the development of efficient and cost-effective methods, for predicting radiative heat transfer rates in practical energy conversion systems.

A number of numerical techniques exist in the literature for the determination of radiative heat transfer rates in various engineering geometries with participating media. Many

new methods. As real problems are almost always of a conjugate nature, according to Howell (1983) and Yang (1986), any practical numerical radiation method should have the following features.

(1) Capability to handle multi-dimensional and complex enclosure geometry.

(2) Good accuracy under all conditions: non-scattering or anisotropic scattering, gray or non-gray, isothermal or non-isothermal participating media.

(3) Ease of application.

(4) Ease of generalization: flexibility in choosing the different order of approximations.

(5) Availability of the accurate intensity field and the integrated quantity.

(6) Computational compatibility with conductive and/or convective heat transfer codes.

(7) Low computation cost.

Therefore, the objective of the research related to the development of radiative heat transfer models for determination of heat transfer rates should be aimed at achieving the above criteria.

The purpose of this review is to acquaint the reader with different approaches in practice presently for solving problems of radiative transport, applicable to heat transfer, in absorbing, emitting, and scattering media. There exists a large body of literature relevant to radiative transport in participating media, and it is not feasible to mention them all in any paper. Also, in these days of numerous journals and allied publications, it is quite possible that some of the major works might have been inadvertently overlooked. Further, any review process is rather an arbitrary process because the authors have to decide what to include, what to omit, and where to start and end. These points, in one sense, reflects the authors' biases, and frankly, the present article is not an exception. In the following sections, reference has only been made to the significant publications related to various methods. A brief introduction of each method has been provided along with its strengths and limitations.

2. Exact analytical method

Analytical methods can be applied to radiative heat transfer problems only in the case of highly idealized situations, like problems with simple geometry and homogeneous participating media having spectrally independent radiative properties. In such situations, exact analytical solutions can be obtained for the integro-differential radiative transfer equation, which is used for determining the variation of intensity in the participating media. RTE applied to one-dimensional plane parallel media have abundant applications in various fields, such as atmospheric sciences, neutron transport etc., where their exact analytical solutions have received much attention (Viskanta & Menguc 1987).

However, exact analytical solutions of RTE for radiative heat transfer in 1-D enclosures and for simple situations are now the subject matter of text books on thermal radiation (see, e.g., Ozisik 1973; Siegel & Howell 1981; Modest 1993). For one-dimensional cylindrical geometry, Heaslet & Warming (1966) have given an exact solution for the case of an

isotropically scattering medium enclosed in an isothermal black cylinder. Azad & Modest (1981) have provided exact analytical solutions for gray, linearly anisotropically scattering media confined in one-dimensional cylindrical enclosures with arbitrary but specified temperature distribution. Loyalka (1969) has also solved a similar kind of problem and his method is also considered exact.

Cheng (1968) has used a rigorous approach to solve the RTE in a two-dimensional rectangular enclosure with absorbing–emitting–isotropic scattering media. Crosbie and his coworkers (Crosbie & Koewing 1979; Crosbie & Dougherty 1980, 1985; Crosbie & Schrenker 1982, 1983, 1985; Crosbie & Farrell 1984; Crosbie & Davidson 1985), have provided exact formulations of RTE for the case of absorbing–emitting–anisotropically scattering media for two- and three-dimensional rectangular and cylindrical enclosures.

For complex radiation problems (for example, problems with inhomogeneous participating media, spectrally dependent radiative properties etc.), the exact analytical solutions of the integro-differential RTE are exceedingly difficult, and explicit solutions are impossible for all but very simple situations. Therefore, for such complicated problems, approximate numerical methods are necessary. These methods have been reviewed in the following sections. Although, exact analytical methods can be adopted only for some simple geometries with many simplifying assumptions, they can be useful in giving quantitative indications for more difficult situations, and serve as benchmarks against which many other approximate numerical methods can be checked.

3. Zonal method

The zonal method, usually known as Hottel's zone method, is one of the most widely used methods for calculating radiation heat transfer in practical engineering systems. It was first developed by Hottel & Cohen (1958) for an absorbing–emitting and non-scattering gray gas with constant absorption coefficients. Hottel & Sarofim (1967), Noble (1975), Smith *et al* (1985) and Nelson *et al* (1986) extended it to deal with non-constant and non-gray absorption coefficients as well as isotropically scattering media.

In this method, the bounding walls and the enclosure volume are divided into a number of zonal elements. Each element is assumed to have a uniform temperature distribution and radiative properties. Then the direct exchange areas, $\overline{s_i s_j}$, between each surface and volume elements are evaluated and the total exchange areas, which include the view factors for the direct exchange as well as those due to multiple reflections, $\overline{S_i S_j}$, are determined. Using these exchange areas, the zonal method expresses the radiant balance of each element in the form:

$$Q_i = Q_{E,i} - \sum_{j}^{N} \overline{S_i S_j}\, E_j. \tag{7}$$

Here the net radiant heat flow, Q_i, from element i, is the difference between emitted radiation, $Q_{E,i}$, and the total radiation received directly and indirectly through the multiple reflection and scattering from all the N elements in the domain. The term E_j is the total blackbody emitted flux, σT_j^4, and $\overline{S_i S_j}$, the total exchange area from element j to i, is a function of the direct exchange area, $\overline{s_i s_j}$, and the radiative properties of the medium.

Once the temperature field of the medium is specified, using the matrix inversion technique, algebraic equation (7) can be solved.

As originally formulated the zonal method has some inherent limitations, such as the treatment of non-gray, temperature-dependent radiative properties (Hottel & Cohen 1958). If the radiative properties are temperature-dependent, the exchange areas have to be re-calculated over and over again, as the temperature field is determined iteratively. For complex geometries, application of the zonal method becomes difficult, as numerous exchange factors between zones must be evaluated and stored in the computer memory. Further, it is often difficult to couple the equations of the zonal method with momentum and general equations for energy conservation, (1), which are usually solved using finite element, finite difference or finite volume techniques. This is mainly because of the different sizes of the required control volumes used in solving radiative equations and momentum and general equations for energy conservation. A coarser grid is desirable for the zonal method. This method can be completely prohibitive if the finer grid scheme as used by the finite difference equations is adopted (Viskanta & Menguc 1987).

Seeing the capability of the zonal method to yield nearly analytical solution, researchers have been trying to enhance its applicability by coupling it with momentum and energy equations. Steward & Tennankore (1979) have coupled the zonal method with finite difference equations in modelling a combustor by adapting two different grid schemes; one for the radiation part and the other for the flow and the temperature field. Smith *et al* (1986) and Sistino (1982) have coupled the zonal method with momentum and energy equations to solve the combined radiative and convective heat transfer in an absorbing–emitting and isotropically scattering medium flowing through a cylindrical duct. More recently, Yuen & Takara (1994) have modified the zonal method to deal with the anisotropic scattering case, but its application to conjugate problems with anisotropic scattering is still awaited.

4. Monte Carlo method

The Monte Carlo method (MCM) is a purely statistical sampling technique. It can exactly simulate almost all important physical processes without much difficulty. In this method, the numerical treatment of the mathematical formulation for incorporation of non-gray, temperature-dependent gas properties and anisotropic scattering situations is easy. Also, the usual difficulties encountered in complex geometries can be circumvented easily. It is because of these advantages that the MCM has been applied to solve many radiative heat transfer problems.

In this method, radiation is assumed to be composed of discrete energy bundles or photons. The computation consists of following the probable path of a finite number of photons, which obey the physical restraints imposed by the RTE, until their final absorption in the system. The path of the photon is determined at each point of emission, absorption–emission, scattering or reflection by a random choice from a set of possible paths. A sufficiently large sample of photons are followed to get statistically stable results.

This method has been used extensively in atmospheric and neutron transport studies. In heat transfer studies, it was first applied to one-dimensional radiative heat transfer

problems in early 1960s by Fleck (1961), Howell & Perlmutter (1964a, 1964b) and Perlmutter & Howell (1964). Radiative heat transfer problems of increasing complexities, such as non-gray, anisotropic scattering media confined in various types of enclosures, that have been investigated by this method have appeared in the works of Taniguchi (1967, 1969), Stockham & Love (1968), Avery *et al* (1969), Gupta *et al* (1983), Abed & Sacadura (1983), Yang *et al* (1983), Tiwari & Liu (1992), Farmer & Howell (1994), and Liu & Tiwari (1994).

Though this method is highly accurate, it is computationally very expensive. To make the Monte Carlo method more attractive for use in conjugate problems, Kobiyama *et al* (1979) and Kobiyama (1986) have studied ways of reducing the computational expense. In certain test problems, they could reduce the computational time by up to 15%. For coupled conduction-radiation heat transfer in semi-transparent media, Abed & Sacadura (1983) have used the finite-difference Monte Carlo method. Liu & Tiwari (1994) have solved conjugate problems for the case of non-gray participating media using narrow band model. Farmer & Howell (1994) have obtained Monte Carlo solutions for radiative heat transfer in a three-dimensional enclosure with anisotropically scattering inhomogeneous media at constant temperature. Most recently, Mishra & Blank (1995) have provided Monte Carlo solutions as a benchmark for two-dimensional enclosures with absorbing–emitting-isotropically scattering media at radiative equilibrium.

Monte Carlo solutions are always very close to the exact analytical solutions, but due to the statistical nature of the method, they fall within a narrow band around the exact analytical solution. Thus due to its accuracy, the method has extensive application to large varieties of engineering problems, and hence it has been concluded that the MCM should be used as a benchmark for comparison of other methods (Viskanta & Menguc 1987; Mishra & Blank 1995). However, like other numerical methods, MCM also has some disadvantages. Its large appetite for computer time, inability to match the required grid size needed for concomitant computation of conduction and/or convection parts of the conjugate problems, and statistical fluctuation of results, limits its use to benchmarking of only purely radiation problems.

5. Multi-flux method

In multi-flux method, to simplify the governing RTE, angular (directional) dependence of intensity is decoupled with its spatial dependence. Intensity is assumed to be uniform over a given interval of the solid angle (Lockwood & Shah 1976; Khalil *et al* 1982). The integro-differential form of RTE is then averaged over each of the finite angular regions to result in a set of flux equations.

Depending on the number of discrete directions over which the solid angle is discretized, different multi-flux methods result. If the solid angle is divided uniformly into two, four or six divisions, the result is two-flux, four-flux and six-flux methods, respectively. The

flux approximations in a three-dimensional cylindrical enclosure containing an absorbing, emitting, and scattering medium.

In general, the accuracy of the flux approximation depends on the choice of the solid angle subdivisions. If there is no intersection between the adjacent subdivisions, more accurate results are expected (Abramzon & Lisin 1984). For rectangular enclosures, the same has also been observed by Siddal & Seluck (1979). If the distribution of radiation intensity is assumed for each subdivision, the general equation given by Abramzon & Lisin (1984) can be simplified and solved simultaneously. If the fluxes in each subdivision are assumed constant, a simpler six-flux model can be obtained from the general flux equations. For an absorbing, emitting and scattering medium in a cylindrical enclosure, Spalding has suggested a simpler six-flux model. A different variation of the six-flux model has also been proposed by Chu & Churchill (1960). Although, Chu and Chirchill's six-flux model is for a one-dimensional plane parallel medium, it is possible to modify it for multi-dimensional enclosures. Starting from this six-flux model, for axi-symmetric cylindrical enclosures, Varma (1979) has obtained a four-flux model.

There are mainly three objections to the multi-flux approximations, for practical engineering problems. Firstly, there could be no coupling between the axial and the radial fluxes, which makes the equations physically unrealistic. Secondly, the approximation of the intensity distribution from which the flux equations are obtained is arbitrary. Lastly, the model equations cannot approximate highly forward scattering correctly, even though, theoretically it is feasible (Viskanta & Menguc 1987).

6. Moment method

In moment method, radiation intensity is expressed by a truncated series representation in terms of the products of angular and spatial functions as

$$i(x, y, z, \theta, \phi) = A_o + \sum_{n=1}^{N} [\zeta^n A_{n,x} + \eta^n A_{n,y} + \mu^n A_{n,z}], \tag{8}$$

where $A^{'}s$ are functions of locations only and ζ, η, and μ are direction cosines in x, y, z directions, respectively (Viskanta & Menguc 1987). Although, this equation is written in Cartesian coordinates, it can be given for any orthogonal system.

This method is the generalization of the Milne–Eddington approximation to the higher order. The Milne–Eddington approximation involves an approximation to the angular dependence of intensity such that the integro-differential RTE is transformed into an ordinary partial differential equation (Modest 1993).

7. Spherical harmonics method

The spherical harmonics method or P_N approximation uses a series approximation to the angularly dependent intensity field to replace the integro-differential RTE with a set

of partial differential equations (Menguc & Viskanta 1985, 1986; Ou & Kuo-Nan 1982; Ratzel & Howell 1983). The intensity field is expanded in terms of the spherical harmonics, where N terms are retained for the P_N approximation.

Like most other approximations, this method was also originally developed for the study of radiative heat transfer in the atmosphere. Latter on, it was modified for solution of neutron transport problems (Davison 1958), and extensively used for one-dimensional heat transfer problems (Kourgnaoff 1952; Chandrasekhar 1960; Ozisik 1973; Cess & Sparrow 1978; Siegel & Howell 1981). Bayazitoglu & Higenyi (1979) have given formulations and solved radiative heat transfer problems for non-scattering Cartesian, cylindrical and spherical media using the first-order P_1 and third-order P_3 spherical harmonics approximations. For an isotropically scattering cylindrical medium, Higeny & Bayazitoglu (1980) and Ahluwalia & Im (1981) have used the P_1 approximation. Ratzel & Howell (1983) have used P_1 and P_3 approximations to solve isotropically scattering two-dimensional rectangular medium problem. Menguc & Viskanta (1985, 1986) have given the general formulation of the P_1 and P_3 approximations for an absorbing, emitting, and anisotropically scattering medium in two-dimensional, finite cylindrical as well as three-dimensional rectangular enclosures. Liu *et al* (1992) have analysed the effect of anisotropic scattering on radiative heat transfer in a two-dimensional rectangular enclosure using P_1 approximation and have compared their results with higher order discrete ordinate method employed by Kim & Lee (1988). It has been shown by Krook (1955) that the moment method and the spherical harmonics method are closely related and completely equivalent (using spherical harmonics, which are functions of direction cosines, and exploiting their orthogonality properties).

The spherical harmonics method or P_N approximation is one of the most tedious and cumbersome of the radiative transfer approximations (Viskanta & Menguc 1987). The formulation of the spherical harmonics method can be very complicated even for the low order P_3 method, in multi-dimensional geometries. Errors can easily result during the long derivations (Menguc & Viskanta 1985, 1986; Ou & Kuo-Nan 1982). Generalization to higher order approximations is not realistic with this method. The P_1 approximation is very accurate, if the optical dimension of the medium is large (i.e., greater than 2) however, it yields inaccurate results for thinner media (Viskanta & Menguc 1987), especially near boundaries. The P_1 approximation becomes less reliable, if the radiation field is anisotropic, i.e., there are large temperature and/or particle concentration gradients in the medium. The P_3 approximation, however, can yield accurate results for an optical dimension as small as 0.5 and for anisotropic radiation fields at the expense of additional computational effort (Menguc & Viskanta 1986; Ratzel & Howell 1983).

8. Discrete ordinate method

The discrete ordinate method or the S_N method was first proposed by Chandrasekhar (1960) in his work on stellar and atmospheric radiation. Originally, this method received little attention in the heat transfer community. Again, like other methods, this method was first systematically applied to problems in neutron transport theory, notably by Lee (1962) and Lathrop (1966). In the beginning, there had been some unoptimized attempts to

apply this method to one-dimensional, planar thermal radiation problems by Love & Grosh (1965), Hsia & Love (1967), Hottel *et al* (1968) and Roux & Smith (1974). However, this method has been applied only very recently to, and optimized for, general radiative heat transfer problems through the pioneering work of Truelove (1987, 1988), Fiveland (1984, 1987,1988), Kim & Lee (1988), Cheong & Song (1995), and Vallion *et al* (1995, 1996).

This method is obtained by discretizing the entire solid angle ($\Omega = 4\pi$) using a finite number of ordinate directions and corresponding weight factors. The N discrete values of direction cosines ζ_n, η_n, μ_n always satisfy the identity $\zeta_n^2 + \eta_n^2 + \mu_n^2 = 1$. The RTE is written for each ordinate direction and the integral terms are replaced by a quadrature summed over each ordinate direction. The ordinate directions are usually the Gaussian or Lobatto quadrature points, although different ordinates and weights have been proposed (Fiveland 1987; O'Dell & Alcouffe 1987). The integral over the angles is then replaced by a quadrature sum. The resulting set of angularly discretized RTE can be spatially discretized by using finite difference method.

This approximation is also not flawless. One of the major drawbacks of this method is the presence of *ray-effects*, which lead to anomalies in scalar flux distribution (Lathrop 1968; Lewis & Miller 1984). These effects are caused by the inability of the discretized intensity distribution to represent the actual continuous distribution fully, especially when radiant energy conservation is not insisted upon. Consequently, radiation can be lost if it does not fall into one of the discrete ordinate directions. Unless an infinite number of directions are used, ray effects cannot be totally eliminated. The ray effects are especially pronounced if there are localized radiation sources in the medium and scattering is less important in comparison to absorption. As the single scattering albedo increases, the radiation field becomes more isotropic and the ray effects become less noticeable. However, with increasing single scattering albedo and/or optical thickness of the medium, the convergence rate may become very slow (Viskanta & Menguc 1987). Also, the computer time and memory requirements of this method are relatively high as compared to P_N approximation and zonal method (Kim 1990).

Another drawback of the discrete ordinate method is the difficulty in extending its application beyond Cartesian geometry (Pomraning 1973). This is evident even in the simple cylindrical system where formulation is complicated by additional restrictions by the presence of curved surfaces (Carlson & Lathrop 1968; Hyde & Truelove 1977).

It should be noted that the basic principle of this method and the multi-flux method are the same, but the difference is in the evaluation of the integral. The flux method uses a trapezoidal rule for the angular integral, whereas the discrete ordinate method uses a higher order Gaussian type quadrature formula, which is more accurate (Khalil *et al* 1982). In the limit of a large number of discrete directions, both methods will converge to the same result.

9. Discrete transfer method

This method, proposed by Shah (1979) and Lockwood & Shah (1981), combines the virtues of the zonal, Monte Carlo, and discrete-ordinate methods. In this method, calculation of radiative heat transfer involves the tracing of representative rays from one surface to another

through the domain of interest. Every bounding control surface in the computational domain is the source of a number of rays chosen to divide the hemisphere about a grid point in a number of equal solid angles. The intensity distribution along each ray is calculated by solving a discretization of the RTE. The RTE along a ray, neglecting scattering can be expressed in the form (Shah 1979)

$$\frac{\mathrm{d}i}{\mathrm{d}s} = -\kappa i + \frac{\kappa \sigma T^4}{\pi} \tag{9}$$

where i is the intensity of radiation, s is the direction along the ray, κ is the absorption coefficient, and σ is the Stephan–Boltzmann constant. For any representative ray, the intensity distribution can be calculated by assuming each control volume as homogeneous. Under this assumption, above equation is integrated to give the recurrence relation,

$$i_{n+1} = i_n \exp(-\kappa \Delta s) + \frac{\kappa \sigma T_c^4}{\pi}(1 - \exp(-\kappa \Delta s)), \tag{10}$$

where i_n and i_{n+1} are the intensity of the ray on entry and exit, respectively, and Δs is the distance travelled in the control volume. T_c is the temperature in the centre of the control volume. Therefore, given the initial intensity at a point on the emitting surface, the change in intensity along the ray can be calculated using (10). The initial intensity is specified by taking the walls to be Lambert surfaces.

For a gray wall, the emitted intensity is dependent on the incident flux q^-

$$q^- = \int_{\phi=2\pi} \int_{\theta=\frac{\pi}{2}} i_w^-(\theta, \phi) \cos\theta \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi, \tag{11}$$

where $i_w^-(\theta, \phi)$ is the incident intensity. In discrete transfer method, the incident flux integral is replaced by a numerical quadrature

$$q^- = \sum i_w^-(\theta, \phi) \cos\theta \sin\theta \Delta\theta \Delta\phi. \tag{12}$$

The values of incident intensity, i_w^-, at a point on a wall are calculated by tracing rays from the point and backtracking from the walls intercepted, applying (10) through each control volume. Hence, for a grey-walled enclosure, the coupling between the incident flux and emitted intensity make the discrete transfer method a guess and correct procedure in which an estimate of the incident flux distribution is iteratively improved.

Lockwood & Shah (1981) showed that very accurate results could be obtained with this method in one- and two-dimensional geometries by increasing the number of rays. More recently, da Graca & Fontes (1993) have solved radiative heat transfer problems in two- and three-dimensional rectangular enclosures with absorbing, emitting, and isotropic scattering media. Their comparison of the results, with other well established methods, such as the zone method, the discrete ordinate method and the P_N method, shows reasonably good agreements.

More recently, Murthy & Choudhary (1992) have employed discrete transfer method to convection-radiation problems in general complex enclosures having non-scattering participating media. Most recently, Cumber (1995) has suggested some improvements to the discrete transfer method and his suggestions have shown some improvements in the

result. However, the improvement in Cumber's (1995) results over that of Shah's (1979) does not seem significant.

Although, this method is claimed to be capable of accounting for anisotropic scattering in the medium, so far no results have been reported or compared against other benchmark methods for scattering media in multi-dimensional enclosures. The results for a one-dimensional scattering medium did not show the same level of agreement with the benchmark results as did the non-scattering medium predictions.

10. Other methods

There are a few methods whose developments have not yet progressed to a level whereby definitive evaluations can be made. However, for the sake of completeness of the review, they are included here.

10.1 *Heat ray method*

This approach to the solution of radiative heat transfer equation for multi-dimensional enclosures has been presented by Taniguchi *et al* (1986) for absorbing–emitting media. This method is based on the Bouguer's law and yields radiant energy absorption distribution in non-isothermal enclosures containing combustion gases. Comparisons of the predictions based on this method with other results show that the method is more accurate and less time-consuming than both zonal and Monte Carlo methods, if the radiative properties such as the absorption coefficient and wall emissivities are constant (Taniguchi *et al* 1986).

10.2 *Discretized intensity method*

This method has been proposed by Shih & Chen (1983) in solving a two-dimensional conduction-radiation problem in absorbing–emitting participating media. It was later used for solving a conjugate convection-radiation recirculating flows problem in a 2-D enclosure (Shih & Ren 1985). Unlike conventional approaches, in which radiative fluxes are assumed to be one-dimensional, this method bypasses the unnecessary use of integro-differential energy transport equation and accounts for two-dimensionality of the radiative fluxes. In its formulation, the conditions at an individual control volume have an explicit effect on the four immediately adjacent control volumes. This method has been shown to be only valid for relatively optically thick participating media (Blank 1992).

10.3 *Finite element approach*

The finite element method has been applied by Razzaque *et al* (1983, 1984), Sokman & Razzaque (1987), Chung & Kim (1984) and Fernandes *et al* (1981), to the combined-mode conduction and/or convection problems with radiation, including scattering and both known temperature and known flux boundary conditions. The finite element method, in principle, gives exact solutions within the errors introduced by the numerical scheme itself, i.e., no approximation is required in the formulation of a problem. According to Howell

Tan (1989) has proposed and applied the product integration method to radiative heat transfer problems. This method, while analogous to zone method or the finite element method, requires evaluation of only three- or two-dimensional integrals for three-dimensional systems. This reduction in the dimension of the required integration, significantly reduces the required computation time. This gain is achieved because for a given degree of approximation of the temperatures (i.e., constant, linear, binomial etc.) within each of the n finite volume or surface elements, the number of calculations in the zonal and finite element method increases as n^2, while in the product integration method, it increases as n (Howell 1988). Although Tan (1989) has incorporated linear anisotropic scattering effect in the analysis of a two-dimensional absorbing, emitting, and scattering media, the formulation is extremely complex.

10.5 *Finite volume approach*

The finite volume approach for radiative heat transfer problems proposed by Raithby & Schneider (1988) and Raithby & Chui (1990) shares the same philosophy as the finite volume technique currently used in predicting fluid flow and convective heat transfer problems. Therefore, it can potentially use the same non-orthogonal computational grid as the fluid flow solver. Conceptually, the method can handle non-uniform radiative properties, anisotropic scattering, and accounts for spectral effects. The application of this method to the radiative heat transfer problems in one- and two-dimensional Cartesian (Raithby & Chui 1990) and axi-symmetric (Chui 1990) enclosures, with absorbing, emitting, and anisotropic scattering media, has shown good comparisons with the existing methods.

10.6 *Collapsed dimension method*

The collapsed dimension method is the latest development which has been proposed by Blank (1994) and Blank & Mishra (1995, 1996). Unlike the rest of the radiation models, it views radiative phenomena quite differently. This method collapses the 3-D radiative information of the problem into its 2-D solution plane. To accomplish this task, the collapsed dimension method makes use of the effective intensity. Each effective intensity contains in itself the information of the conglomeration of all the conventional intensities, focused on the same point and contained in a plane which is both inclined at the same angle as the effective intensity and also normal to the solution plane. This novel procedure thus eliminates the use of the solid angles and reduces both the complexity and computational expense involved in the solution of such problems. This method has been shown to give near analytical results for all ranges of optical thicknesses, including optically very thin cases where most of the existing methods fail to perform well. Since this method views all radiative phenomena in the 2-D solution plane, the computational time taken by this method for solution of any particular problem will be the least. For real life conjugate

problems, this method seems to be one of the best. However, this method has not yet been tested for such complicated situations.

11. Conclusions

Based on the above discussions, it is evident that no single radiation model can be regarded as the best model for all types of problems. Each method has its own merits and demerits. Thus, to ascertain whether a particular method is appropriate for a given problem, comparisons have to be made against the benchmark results obtained either from experiments or exact analytical solutions. Although, the Monte Carlo method's applicability to conjugate problems is very restricted, for the majority problems dealing purely with radiative transport, its solutions serve as benchmark results. The zonal method solution is also considered exact and has been successfully applied to a large class of problems. The P_N method and the discrete ordinate method, even being approximate methods, have the most diverse applications, in all types of simple and complex problems (conjugate problems). However, these methods do not perform well at low optical thickness situations. The discrete transfer method has been found to give good results for different types of enclosures with simple boundary conditions and simple participating media conditions but for complicated situations it has not been tested. Finite element and finite volume approaches to radiative transport problems have the added advantage in conjugate problems in which most of the methods have difficulties in matching the grid sizes of radiative part of the problem with conductive and/or convective part. The collapsed dimension method has been found to yield very accurate results for the problems tested so far and is believed to take least time compared to the existing methods. Unlike, the P_N and discrete ordinate methods, this method works extremely well at very low optical thickness conditions. Moreover, this method has not been tested so far for a large class of problems and thus no conclusive remarks can be made about complex situations. In conclusion, depending upon the computing facility and desired accuracy of the result etc. for any problem, it is the user's discretion to adopt a particular method.

List of symbols

c_v specific heat at constant volume, [W/m^2 kg K];
E emissive power, [W/m^2];
E_b blackbody emissive power, [W/m^2];
G incident radiant energy, [W/m^2];
i intensity, [W/m^2];
i_b Planck's blackbody intensity, [W/m^2];
k thermal conductivity, [W/m K];
N number of volume elements in zonal method;
p phase function;
Q energy flow rate, [W];

s geometric path, [m];
$\overline{s_i s_j}$ direct exchange areas in zonal method, [m^2];
$\overline{S_i S_j}$ total exchange areas in zonal method, [m^2];
T temperature, [K];
t time, [s];
$\mathbf{v}$ velocity vector, [m/s];
x, y, z Cartesian coordinates, [m];
η direction cosine, [-];
θ zenith angle, [rad];
τ optical thickness, optical coordinate, [-];
τ_L enclosure optical thickness, [-];
κ absorption coefficient, [m^{-1}];
μ direction cosine, [-];
μ dynamic viscosity, [kg/m s];
ζ direction cosine, [-];
σ Stephan–Boltzmann constant $= 5.670 \times 10^{-8}$ W/m^2K^4;
σ_s scattering coefficient, [m^{-1}];
ϕ azimuthal angle, [rad];
Φ dissipation function, [J/kg m^2];
ω single scattering albedo, [-];
ρ density, [g/m^3].

Subscripts

a absorption;
b blackbody;
s scattering;
λ wavelength dependent.

Superscripts

$+$ positive direction;
$-$ negative direction;
$'$ dummy variables.

References

Abed A Al, Sacadura J F 1983 A Monte Carlo-finite difference method for coupled radiation-conduction heat transfer in semitransparent media. *Trans. ASME J. Heat Transfer* 105: 931–933

Abramzon M N, Lisin F N 1984 Methods for solving the radiant-transfer equations in cylindrical geometry. *High Temperature* 22: 95–100

Ahluwalia R K, Im K H 1981 Combined conduction, convection, gas radiation and particle radiation in MHD diffusers. *Int. J. Heat Mass Transfer* 24: 1421–1430

Avery L W, House L L, Skumanchi A 1969 Radiative transport in finite homogeneous cylinders by the Monte Carlo technique. *J. Quant. Spectrosc. Radiat. Transfer* 9: 519–531

Azad F H, Modest M F 1981 Evaluation of the radiative heat flux in absorbing, emitting and linear-anisotropic scattering cylindrical media. *Trans. ASME J. Heat Transfer* 103: 350–356

Bayazitoglu Y, Higenyi J 1979 The higher-order differential equations of radiative transfer: P_3 approximation. *AIAA J.* 17: 424–431

Blank D A 1992 Modified, discretized-intensity based, split radiation calculation procedure for use in full simulation studies of combustion in axi-symmetric piston engines. *Numer. Heat Transfer* A22: 199–222

Blank D A 1994 The cartesian collapsed-dimension method for use in numerical 2-D radiative calculations in absorbing–emitting media. *Int. J. Numer. Methods Eng.* 37: 3023–3036

Blank D A, Mishra S C 1995 Use of the 2-D collapsed dimension method in absorbing–emitting media with isotropic scattering. In *Proceedings of the First International Symposium on Radiative Heat Transfer* (New York: Begell House) pp 138–151

Blank D A, Mishra S C 1996 Use of the 2-D collapsed dimension method in gray enclosures with absorbing-emitting–isotropic scattering media in radiative equilibrium. *Numer. Heat Transfer* B30: 469–481

Carlson B G, Lathrop K D 1968 Transport theory – the method of discrete ordinates. In *Computing methods in reactor physics* (eds) H Greenspan, C N Kebler, D Okrent (New York: Gordon & Breach) pp 269–308

Cess R D, Sparrow E M 1978 *Radiative heat transfer* Augmented edn. (Washington, DC: Hemisphere)

Chandrasekhar S 1960 *Radiative transfer* (New York: Dover)

Cheng P 1968 Thermal control and radiation. In *Progress in astronautics and aeronautics* (ed.) C L Tien (New York: Gordon and Breach) p 171

Cheong K-B, Song T-H 1995 Examination of solution methods for the second-order discrete ordinate formulation. *Numer. Heat Transfer* B27: 155–173

Chu C M, Churchill S W 1960 Numerical solution of problems in multiple scattering of electromagnetic radiation. *J. Phys. Chem.* 59: 855–863

Chui E H-K 1990 *Modeling of radiative heat transfer in participating media by the finite volume method.* Ph D thesis, University of Waterloo, Ontario

Chung T J, Kim J Y 1984 Two-dimensional, combined-mode heat transfer by conduction, convection, and radiation in emitting, absorbing, and scattering media-solution by finite elements. *Trans. ASME J. Heat Transfer* 106: 448–452

Crosbie A L, Davidson G W 1985 Dirac-delta function approximations to the scattering phase function. *J. Quant. Spectrosc. Radiat. Transfer* 33: 391–409

Crosbie A L, Dougherty R L 1980 Two-dimensional radiative transfer in a cylindrical geometry with anisotropic scattering. *J. Quant. Spectrosc. Radiat. Transfer* 25: 551–569

Crosbie A L, Farrell J B 1984 Exact formulation of multiple scattering in a three-dimensional cylindrical geometry. *J. Quant. Spectrosc. Radiat. Transfer* 31: 397–416

Crosbie A L, Koewing J W 1979 Two-dimensional radiative transfer in a finite scattering planar medium. *J. Quant. Spectrosc. Radiat. Transfer* 21: 573–595

Crosbie A L, Schrenker R G 1982 Exact expressions for a radiative transfer in a three-dimensional rectangular geometry. *J. Quant. Spectrosc. Radiat. Transfer* 28: 507–526

Crosbie A L, Schrenker R G 1983 Multiple scattering in a two-dimensional rectangular medium exposed to collimated radiation. *J. Quant. Spectrosc. Radiat. Transfer* 33: 101–125

Crosbie A L, Dougherty R L 1985 Two-dimensional linearly anisotropic scattering in a finite thick cylindrical medium exposed to a laser beam. *J. Quant. Spectrosc. Radiat. Transfer* 33: 487–520

Crosbie A L, Schrenker R G 1985 Multiple scattering in a two-dimensional rectangular medium exposed to collimated radiation. *J. Quant. Spectrosc. Radiat. Transfer* 33: 101–125

Cumber P S 1995 Improvements to the discrete transfer method of calculating radiative heat transfer. *Int. J. Heat Mass Transfer* 38: 2251–2258

da Graca C, Fontes F T 1993 Multidimensional modeling of radiative heat transfer in scattering media. *Trans. ASME J. Heat Transfer* 115: 486–489

Davison B 1958 *Neutron transport theory* (Oxford: Clarendon)

Farmer J T, Howell J R 1994 Monte Carlo prediction of radiative heat transfer in inhomogeneous, anisotropic, nongray media. *J. Thermophys. Heat Transfer* 8: 133–139

Fernandes R, Francis J, Reddy J N 1981 A finite element approach to combined conductive and radiation heat transfer in a planar medium. In *Heat transfer and thermal control. Progress in aeronautics and astronautics* (New York: AIAA) pp 93–109

Fiveland W A 1984 Discrete-ordinate solutions of the radiative transport equation for rectangular enclosures. *Trans. ASME J. Heat Transfer* 106: 699–706

Fiveland W A 1987 Discrete ordinate methods for radiative heat transfer in isotropically and anisotropically scattering media. *Trans. ASME J. Heat Transfer* 109: 809–812

Fiveland W A 1988 Three-dimensional radiative heat transfer solutions by the discrete ordinate method. *J. Thermophys. Heat Transfer* 2: 309–316

Fleck J A 1961 The calculation of nonlinear radiation transport by Monte Carlo method: Statistical physics. *Methods Comput. Phys.* 1: 43–65

Gupta R P, Wall T F, Truelove J S 1983 Radiative scatter by fly ash in pulverized-coal-fired furnaces: Application of the Monte Carlo method to anisotropic scattering. *Int. J. Heat Mass Transfer* 26: 1649–1660

Heaslet M A, Warming R F 1966 Theoretical predictions of radiative heat transfer in homogeneous cylindrical medium. *J. Quant. Spectrosc. Radiat. Transfer* 6: 751–774

Higenyi J, Bayazitoglu Y 1980a Differential approximation of radiative heat transfer in a gray medium. *Trans. ASME J. Heat Transfer* 102: 719–723

Higenyi J, Bayazitoglu Y 1980b Radiative transfer of energy in a cylindrical enclosure with heat generation. *AIAA J.* 18: 723–726

Hottel H C, Cohen H S 1958 Radiant heat exchange in a gas filled enclosure: Allowance for nonuniformity of gas temperature. *AIChE J.* 4: 3–14

Hottel H C, Sarofim A F 1967 *Radiative heat transfer* (New York: McGraw-Hill)

Hottel H C, Sarofim A F, Evans L B, Vasalos I A 1968 Radiative transfer in anisotropically scattering media: Allowance for fresnel reflection at the boundaries. *Trans. ASME J. Heat Transfer* 90: 56–62

Howell J R 1983 Radiative transfer in multi-dimensional enclosures with participating media. ASME, paper no. 83-HT-32

Howell J R 1988 Thermal radiation in participating media: The past, the present, and some possible future. *Trans. ASME J. Heat Transfer* 110: 1220–1229

Howell J R, Perlmutter M 1964a Monte Carlo solution of radiant heat transfer in a nongray, nonisothermal gas with temperature-dependent properties. *AIChE J.* 10: 562–567

Howell J R, Perlmutter M 1964b Monte Carlo solution of thermal transfer through radiant media between gray walls. *Trans. ASME J. Heat Transfer* 86: 116–122

Hsia H M, Love T J 1967 Radiative heat transfer between parallel plates separated by a nonisothermal medium with anisotropic scattering. *Trans. ASME J. Heat Transfer* 89: 197–204

Hyde D J, Truelove J S 1977 The discrete ordinate approximation for radiative heat transfer. Technical Report AERE-R 8502

Khalil H, Shultis J K, Lester T W 1982 Comparison of three numerical methods for evaluation of radiant energy transfer in scattering and heat generating media. *Numer. Heat Transfer* 5: 235–252

Kim T-K 1990 *Radiation and combined mode heat transfer analysis in absorbing, emitting, and Mie-anisotropic scattering media using S-N discrete ordinate method.* Ph D thesis, University of Minnesota, Minneapolis, MN

Kim T-K, Lee H 1988 Effect of anisotropic scattering on radiative heat transfer in two-dimensional rectangular enclosures. *Int. J. Heat Mass Transfer* 31: 1711–1721

Kobiyama M 1986 A study on the reduction of computing time of the Monte Carlo method applied to the radiative transfer. *Bull. Jap. Soc. Mech. Eng.* 29: 3000–3006

Kobiyama M, Taniguchi H, Saito T 1979 The numerical analysis of heat transfer combined with radiation and convection (1st report, the effect of two-dimensional radiative transfer between isothermal parallel plates). *Bull. Jap. Soc. Mech. Eng.* 22: 707–714

Kourgnaoff V 1952 *Basic methods in transfer problems.* (London: Oxford University Press)

Krook M 1955 On the solution of equation of transfer, I. *Astrophys. J.* 122: 488–497

Kuo K 1986 *Principles of combustion* (New York: Wiley) chap. 7

Lathrop K D 1966 Use of discrete ordinate methods for solution of photon transport problems. *Nuclear Sci. Eng.* 24: 381–388

Lathrop K D 1968 Ray effects in discrete ordinate equations. *Nuclear Sci. Eng.* 32: 357–369

Lee C E 1962 The discrete S_n approximation to transport theory. Technical Report Technical Information Series Report LA2595, Lawrence Livermore Laboratory.

Lewis E E, Miller W F Jr 1984 *Computational methods of neutron transport* (New York: John Wiley and Son)

Liu F, Garbett G S, Swithenbank J 1992 Effect of anisotropic scattering on radiative heat transfer using the p_1-approximation. *Int. J. Heat Mass Transfer* 35: 2491–2499

Liu J, Tiwari S N 1994 Radiative interactions in chemically reacting compressible nozzle flows using Monte Carlo simulation. In: *6th AIAA/ASME Joint Thermophysics and Heat Transfer Conference*, Colorado Springs, CO (AIAA) pp 1–13

Lockwood F C, Shah N G 1976 An improved flux model for the calculation of radiation heat transfer in combustion chambers. In: *Proceedings of the 16th National Heat Transfer Conference*, St. Louis (ASME) pp 2–7

Lockwood F C, Shah N G 1981 A new radiation solution method for incorporation in general combustion prediction procedures. In *Proceedings of the Eighteenth International Symposium on Combustion* (Pittsburg, PA: The Combustion Institute) pp 1405–1414

Love T J, Grosh R J 1965 Radiative heat transfer in absorbing, emitting, and scattering media. *Trans. ASME J. Heat Transfer* 87: 161–166

Loyalka S K 1969 Radiative heat transfer between parallel plates and concentric cylinders. *Int. J. Heat Mass Transfer* 12: 1513–1517

Manickavasagam S, Menguc M P 1993 Effective optical properties of pulverized coal particles determined from FT-IR spectrometer experiments. *Energy Fuels* 7: 860–869

Menguc M P, Viskanta R 1985 Radiative transfer in three-dimensional rectangular enclosures containing inhomogeneous, anisotropically scattering media. *J. Quant. Spectrosc. Radiat. Transfer* 33: 533–549

Menguc M P, Viskanta R 1986 Radiative transfer in axisymmetric, finite cylindrical enclosures. *Trans. ASME J. Heat Transfer* 108: 271–276

Mishra S C, Blank D A 1995 High to low optical thickness Monte Carlo solutions of the radiative

Modest M F 1993 *Radiative heat transfer* (Singapore: McGraw-Hill)
Murthy J Y, Choudhary D 1992 Computation of participating radiation in complex geometries. In *28th National Heat Transfer Conference*, San Diego, CA (ASME, Heat Transfer Div.) pp 153–160
Nelson H F, Look D C, Crosbie A L 1986 Two-dimensional radiative back-scattering from optically thick media. *Trans. ASME J. Heat Transfer* 108: 619–625
Noble J J 1975 The zone method: Explicit matrix relations for total exchange areas. *Int. J. Heat Mass Transfer* 18: 261–269
O'Dell R D, Alcouffe R E 1987 Transport calculations for nuclear analysis: Theory and guidelines for effective use of transport codes. Technical Report La-10983-MS, Los Alamos National Laboratory
Ou S-C S, Kuo-Nan L 1982 Generalization of the spherical harmonics method to radiative transfer in multi-dimensional space. *J. Quant. Spectrosc. Radiat. Transfer* 28: 271–288
Ozisik M N 1973 *Radiative transfer and interactions with conduction and convection* (New York: John Wiley & Sons)
Perlmutter M, Howell J R 1964 Radiant heat transfer through a gray gas between concentric cylinders using Monte Carlo. *Trans. ASME J. Heat Transfer* 86: 169–179
Pomraning G C 1973 *The equations of radiation hydrodynamics* (Oxford: Pergamon)
Raithby G D, Chui E H 1990 A finite volume method for predicting a radiant heat transfer in enclosures with participating media. *Trans. ASME J. Heat Transfer* 112: 415–423
Raithby G D, Schneider G E 1988 In *Handbook of numerical heat transfer* (New York: John Wiley & Sons) pp 241–291
Ratzel A C III, Howell J R 1983 Two-dimensional radiation in absorbing–emitting media using the P_N approximation. *Trans. ASME J. Heat Transfer* 105: 333–340
Razzaque M M, Klein D E, Howell J R 1983 Finite element solution of radiative heat transfer in a two-dimensional rectangular enclosures with gray participating media. *Trans. ASME J. Heat Transfer* 105: 933–936
Razzaque M M, Howell J R, Klein D E 1984 Coupled radiative and conductive heat transfer in a two-dimensional enclosure with gray participating media using finite elements. *Trans. ASME J. Heat Transfer* 106: 613–619
Roux J A, Smith A M 1974 Radiative transport analysis for plane geometry with isotropic scattering and arbitrary temperature. *AIAA J.* 12: 1273–1277
Shah N G 1979 *New method of computation of radiation heat transfer combustion chambers.* Ph D thesis, Imperial College, University of London, London
Shih T M, Chen Y N 1983 A discretized-intensity method proposed for two-dimensional systems enclosing radiative and conductive media. *Numer. Heat Transfer* 6: 117–134
Shih T M, Ren A L 1985 Combined convective and radiative recirculating flows in enclosures. *Numer. Heat Transfer* 8: 149–167
Siddal R G, Seluck N 1979 Evaluation of a new six-flux model for radiative transfer in rectangular enclosures. *Trans. IChE.* 7: 163–169
Siegel R, Howell J R 1981 *Radiative heat transfer* 2nd edn (Washington, DC: Hemisphere)
Sistino A J 1982 Mean beam length and zone method (without and with scattering) for a cylindrical enclosure. ASME paper no. 82-HT-3
Smith T F, Shen Z F, Al-Turki A M 1985 Radiative and conductive transfer in a cylindrical enclosure for a real gas. *Trans. ASME J. Heat Transfer* 104: 482–485
Smith T F, Byun K H, Ford M J 1986 *Heat transfer* (Washington, DC: Hemisphere) vol. 2, pp 803–808
Sokman C N, Razzaque M M 1987 Finite element analysis of conduction radiation heat transfer

in an enclosure with heat flux boundary conditions. In *Radiation, phase change heat transfer, and thermal systems* (ASME Heat Transfer Div.) vol. 81, pp 17–23

Steward F R, Tennankore K N 1979 *J. Inst. Energy* 53: 107–112

Stockham L W, Love T M 1968 Radiative heat transfer from a cylindrical cloud of particles. *AIAA J.* 6: 1935–1940

Tan Z 1989 Radiative heat transfer in multidimensional emitting, absorbing, and anisotropic scattering media – mathematical formulation and numerical method. *Trans. ASME J. Heat Transfer* 111: 141–147

Taniguchi H 1967 Temperature distribution of radiant gas calculated by Monte Carlo method. *Bull. Jap. Soc. Mech. Eng.* 10: 975–988

Taniguchi H 1969 The radiative heat transfer of gas in a three dimensional system calculated by Monte Carlo method. *Bull. Jap. Soc. Mech. Eng.* 12: 67–78

Taniguchi H, Yang W J, Kudo K, Hayasaka H, Oguma M, Kusama M, Nakamachi I, Okigani N 1986 *Heat transfer* (Washington, DC: Hemisphere) vol. 2, pp 757–762

Tien C L 1988 Thermal radiation in packed and fluidized beds. *Trans. ASME J. Heat Transfer* 110: 1230–1242

Tiwari S N, Liu J 1992 Investigation of radiative interaction in laminar flows of nongray gases using Monte Carlo simulation. In *Proceedings of the National Heat Transfer Conference*, San Diego, CA (ASME) pp 187–195

Truelove J S 1987 Discrete-ordinate solutions of the radiative transport equation. *Trans. ASME J. Heat Transfer* 109: 1048–1051

Truelove J S 1988 Three-dimensional radiation in absorbing–emitting-scattering media using the discrete-ordinate approximation. *J. Quant. Spectrosc. Radiat. Transfer* 39: 27–31

Vaillon R, Lallemand M, Lemonnier D 1995 Radiative equilibrium in axisymmetric semi-transparent gray shells using the discrete ordinate method. In *Proceedings of the First International Symposium on Radiative Heat Transfer*, Kusadasi, Turkey

Vaillon R, Lallemand M, Lemonnier D 1996a *EUROTHERM* (eds) D Lemonnier, J B Savlnier, M Fiebig (Amsterdam: Elsevier) pp 1994–2000

Vaillon R, Lallemand M, Lemonnier D 1996b Radiative heat transfer in orthogonal curvilinear coordinates using the discrete ordinate method. *J. Quant. Spectrosc. Radiat. Transfer* 55: 7–17

Varma S A 1979 *Pulverized coal combustion and gassification* (New York: Plenum) pp 311–315

Viskanta R, Menguc M P 1987 Radiation heat transfer in combustion systems. *Prog. Energy Combust. Sci.* 13: 97–160

Yang K T 1986 Numerical modeling of natural convection-radiation interactions in enclosures. In *Proceedings of the 8th International Heat Transfer Conference* (Washington, DC: Hemisphere) vol. 1, pp 131–140

Yang Y S, Howell J R, Klein D E 1983 Radiative heat transfer through a randomly packed bed of spheres by the Monte carlo method. *Trans. ASME J. Heat Transfer* 105: 325–332

Yuen W W, Takara E E 1994 Development of a generalized zonal method for analysis of radiative transfer in absorbing and anisotropically scattering media. *Numer. Heat Transfer* B25: 75–96

Vol. 23, Part 2, April 1998

SĀDHANĀ

Academy Proceedings in Engineering Sciences

CONTENTS

Indexed in CURRENT CONTENTS

ISSN 0256-2499

Edited and published by N Mukunda for the Indian Academy of Sciences, Bangalore 560 080.
Typeset and Printed at Thomson Press (India) Ltd., Faridabad 121 007.